Rationing Medicine Through Bureaucracy: Authorization Restrictions in Medicare

By Zarek Brot-Goldberg, Samantha Burn, Timothy Layton, and Boris Vabson*

High administrative costs in U.S. health care have provoked worry among policymakers, but much of these costs are generated by managed care policies that trade off bureaucratic sludge against reductions in moral hazard. We study this trade-off for prior authorization restriction policies in Medicare Part D, where low-income beneficiaries are randomly assigned to default plans. Beneficiaries who face restrictions on a drug reduce their use by 23%. Roughly half of patients are diverted to another related drug, while the other half are diverted to no drug. These policies generated net fiscal savings, reducing drug spending by $61 per patient-year, while creating $7.30 in paperwork costs. Revealed preference approaches suggest that the foregone drug value is below the cost savings.

Extremely preliminary. Please do not cite or circulate.

* We thank Frank Gaunt for excellent research assistance, and Jason Abaluck, Giovanni Compiani, Stuart Craig, Josh Gottlieb, Bruce Landon, Lee Lockwood, Anna Zink, and seminar participants at APPAM 2021, Chicago Booth, Chicago Health Economics Workshop, IIOC 2022, Notre Dame, NYU, UC Berkeley, the 2021 NBER Public/Insurance Joint Meetings, and the 12th Annual Empirical Health Law Conference for helpful comments. We thank Arnold Ventures and the Becker-Friedman Institute for their financial support of this work. All errors are our own.

The cost of bureaucratic administration makes up a substantial portion of the costs of providing health care in the U.S., estimated between 20 and 34% of health care expenditures (Woolhandler et al., 2003; Dunn et al., 2020; Himmelstein et al., 2020), roughly 1-4% of GDP. The focus of most academic and policy discussion regarding this bureaucratic system has framed it as generating substantial waste, arguing that it causes the U.S. health care system to be "on a production possibility frontier that is interior to that of other countries" (Cutler and Ly, 2011). The promise of reducing these costs is a key component of arguments supporting U.S. health care reform.[1]

However, roughly half of administrative costs are spent on policies that aim to rein in costs: Auditing claims for fraud, overbilling, or wasteful care; and compliance with managed care policies that restrict access to costly providers, services, and drugs (Cutler, 2020). These activities explicitly exist to regulate patient and provider behavior and reduce health care spending. While there are potential cost reductions from making existing administration more efficient (Cutler et al., 2012), reducing bureaucracy by significant magnitudes would require reducing these activities as well.

In this paper, we take seriously the idea that bureaucracy has both costs and benefits. Bureaucratic systems trade off administrative burden for potential reductions in moral hazard and lower costs of insurance provision. We characterize this trade-off for a specific case: Prior authorization restrictions for prescription drugs. Under such policies, patients can only receive insurance coverage for certain drugs (typically high-cost, on-patent drugs) if they receive explicit authorization; otherwise they must pay the full cost out of pocket. Getting authorization requires their prescribing physician to fill out pre-specified paperwork making the case for why the patient should receive the drug. The goal of these policies is to restrict access to costly drugs to only those patients who have the highest value. However, prior authorization comes with costs: Making authorization requests is a major source of administrative effort, requiring an average of 20.4 manpower hours per physician per week for physician practices, their second greatest administrative burden behind billing (Casalino et al., 2009).

We conceptualize prior authorization as an 'ordeal,' in the terminology of Nichols and Zeckhauser (1982). Under such a mechanism, physicians reveal their beliefs about their patients' valuations for the restricted drugs through their willingness to exert costly effort to do paperwork on the patients' behalf, allowing payers to ration payments to their highest-value uses. Such mechanisms can therefore generate waste even when applied optimally. The welfare effects of the use of prior authorization contrast the bureaucratic 'sludge' costs applied on inframarginal patients who go through the authorization process against the reductions in moral hazard for marginal patients who are deterred; they therefore

[1]For example, one argument for single-payer reform is that traditional Medicare spends less per beneficiary in administration than private insurers (Archer, 2011; Frakt, 2018). Administrative costs are also central to explaining differences in costs of health care provision between the U.S. and other OECD countries (Pozen and Cutler, 2010).

depend on the size and composition of these marginal and inframarginal groups.

We study prior authorization empirically in Medicare Part D, the public drug insurance program for the elderly. We focus on the Low-Income Subsidy (LIS) program. This program has two important features that make it useful for studying the effects of prior authorization. First, there is effectively no cost-sharing for LIS beneficiaries. Instead, the *only* significant factors affecting drug demand in this population are whether a drug faces a prior authorization restriction or is outright excluded from coverage. Second, the default rules in the LIS program often assign beneficiaries to a randomly-chosen plan, and default compliance is substantial, providing exogenous variation in exposure to prior authorization at the person-drug level. Since rationing with cost-sharing is impossible for this population, the use of prior authorization restrictions are substantial: In 2015, prior authorization policies covered roughly 5% of prescriptions making up 22% of drug spending.

We start by measuring the effect of prior authorization on drug utilization. Our research design compares, *within a given drug*, utilization for beneficiaries who are randomly assigned to plans that have authorization restrictions on that drug against those assigned to plans that cover the drug without restriction, while controlling for secular effects of plans on overall utilization. We find that prior authorization restrictions reduce the use of focal drugs by 23.6%, with larger relative effects among non-white, sicker patients. To understand what deterred patients do, we estimate a discrete-choice model of demand for prescription drugs within therapeutic classes. We impose a nested logit structure, allowing the nesting parameter to govern the extent of substitution on the intensive margin, to other drugs, vs. the extensive margin, to no drug at all. We estimate that roughly half of patients substitute on each of the two margins.

We then use our estimated demand system to compare overall utilization and spending under the status quo against a counterfactual world in which no authorization restrictions were ever imposed. We find that, accounting for substitution, prior authorization restrictions reduced total drug spending spending by 3.1%, or approximately $61 per beneficiary-year.

While prior authorization lowers the cost of insurance provision, it also generates costs from the administrative burden it imposes on providers. Our data does not permit us to directly measure these costs, so we calibrate per-application accounting costs and combine them with our demand system to estimate the size of the burden. Our preferred calibration calculates that for each dollar saved by prior authorization, it induces $0.12 of administrative costs, with a total burden of $7.30 per beneficiary-year. While the costs of bureaucracy are large, they are second-order relative to the effects on utilization, and therefore prior authorization *reduces* total spending on health care.

These calibrations suggest that the administrative burden imposed by managed care policies like prior authorization is second-order relative to the induced program savings. The question remaining is whether moral hazard is large enough

to justify such policies; that is, whether the lost patient value of deterred drugs is below the fiscal savings recouped by not paying for them. In a typical product market, an econometrician would use the estimated demand system and revealed preference to estimate private value. In this setting, doing so is impossible given the absence of (consumer-facing) prices and the presence of an agent (the prescribing physician) who may exert their own preferences.

We take two approaches. First, we attempt to quantify the effect of prior authorzation on patient health, both overall for all drug classes, as well as specifically for patients taking oral anticoagulants, a drug class where poor adherence can lead to immediate health effects. Both of these exercises produce noisy results that are indistinguishable from zero, although we can reject large health effects.

Second, we use a revealed preference approach to estimate forgone consumer surplus from reduced consumption of focal drugs. Since we do not observe prices in our main sample, we leverage a different sample, of beneficiaries who transition from the unsubsidized component of Medicare Part D into the LIS program. This transition shifts out-of-pocket prices from a positive amount to approximately zero. We use this to estimate a price semi-elasticity of demand of roughly 0.38. We then use this elasticity to back out implied consumer surplus under various assumptions regarding which part of the demand curve is screened out by prior authorization. Our estimates range from consumer surplus losses of \$2.88 if prior authorization screens out the lowest-value beneficiaries, to \$21.05 if it screens out a random subset of beneficiaries, to \$50.23 if it screens out the highest-value beneficiaries. We also recover analogues using physicians' implied administrative cost semi-elasticity of demand coming from our model of drug demand; under the assumption that physicians are perfect agents for their patients, this approach produces much smaller consumer surplus losses under our screening assumptions, at \$0.93, \$8.75, and \$20.98, respectively. All of these estimates are still below the net fiscal savings generated by the restrictions.

These results lead to a two-handed conclusion for the welfare effects of bureaucratic restrictions like prior authorization. On the one hand, they have some undesirable targeting properties; inducing larger effects among sicker and nonwhite populations, as well as reducing drug use overall rather than merely rerouting it all to cheaper generics. This conclusion is consistent with recent scholarship in economics and public administration finding that administrative barriers to social program take-up tend to screen vulnerable populations out rather than in (Heinrich, 2016; Herd and Moynihan, 2018; Deshpande and Li, 2019; Finkelstein and Notowidigdo, 2019; Homonoff and Somerville, forthcoming; Shepard and Wagner, 2021), although this scholarship has studied beneficiary-facing barriers, rather than ones that are intermediated through service providers.[2]

[2]This literature has also tended to consider only the effect on the margin, rather than the burden placed on the inframarginal recipients. This is sensible given that a change which *reduces* targeting efficiency while also raising inframarginal burdens must be inefficient no matter how large the burden it imposes is. A smaller literature has considered how to design in-kind transfer programs trading off targeting efficiency against reduced value for inframarginal recipients, although that literature has focused

On the other hand, our calibrations suggest that the program savings are reasonably in excess of the costs of these policies, suggesting the presence of moral hazard, necessitating stricter rationing. Our results, that bureaucratic review may generate positive social welfare effects, are in line with recent work on claims audits by Eliason et al. (2021) and Shi (2021) and on opioid monitoring by Alpert et al. (2020).[3] Moreover, other rationing mechanisms frequently used in health care, such as price rationing through the use of deductibles, have been documented to have even worse targeting properties (Sood et al., 2013; Baicker et al., 2015; Brot-Goldberg et al., 2017). Bureaucratic review burdens on relatively-better-informed physicians may be the superior mechanism for allocating drugs and other medical services even if they fail to achieve the second-best outcome. In general, rationing mechanisms targeting providers may be more efficient, consistent with positive results on provider price rationing from Ho and Pakes (2014) and Song et al. (2019).

Whether bureaucracy in health care is an overall force for good or for bad, our results suggest that its *effects on output* are of greater orders of magnitude than the direct costs of its operation. This is critical since scholarship has largely focused on accounting and measurement (Casalino et al., 2009; Cutler et al., 2012; Gottlieb et al., 2018; Dunn et al., 2020, 2021), whereas we know less about how bureaucracy affects prices and quantities, both in health care and beyond.[4]

The paper proceeds as follows. In Section I, we describe how prior authorization is used in practice and sketch a simple conceptual model that highlights what quantities need to be estimated to think about the welfare effects of such policies. In Section II, we describe Medicare Part D, the LIS program, and the data we use. We also provide descriptive statistics on the use of prior authorization policies in this setting. In Section III, we estimate reduced-form regressions of drug use on its prior authorization restriction status. In Section IV, we estimate a structural model of drug demand with the goal of characterizing intensive vs. extensive margin substitution patterns. In Section V, we use the results of this model to measure the trade-off between spending reductions and paperwork costs using calibrated measures of provider costs. In Section VI, we measure effects

on the form of provision rather than the explicit rationing mechanism; see, e.g., Lieber and Lockwood (2019) and Waldinger (2021).

[3]There is also a small existing literature on the quantity effects of prior authorization policies (Seabury et al., 2014; Sarig, 2020). Closest to our work is Dillender (2018), who also documents poor targeting effects of prior authorization for a small set of abuse-prone drugs in the Texas worker's compensation insurance system. This literature has generally used time-series variation in prior authorization restrictions, which may be confounded by the effects of evolving patterns of drug utilization. This evolution is especially important among frequently-restricted drugs, which tend to be new.

[4]We note three important exceptions. First, Bajari et al. (2014) show that the possibility of future bureaucratic hold-ups lowers bids in government procurement auctions. Second, Gennaioli et al. (2020) show that higher claim denial frequency in homeowner insurance lowers equilibrium premiums. Third, Dunn et al. (2021) show that claim denial frequency in Medicaid affects the willingness of physicians to accept Medicaid patients. Our contribution is closest to Dunn et al. in that we consider effects on quantities, but we study quantity decisions *within* an existing relationship rather than the bureaucracy's effect on relationship formation; since Medicare Part D insurers do not directly contract with providers, such a channel is shut down in our setting.

of authorization restrictions across demographics and on patient health to assess the efficiency of conserved care. Section VII concludes.

I. Prior Authorization Restrictions in Theory and Practice

A. Prior Authorization Restrictions in Practice

The vast majority of health insurance is provided by managed care organizations (MCOs), private insurers who provide insurance coverage, but place restrictions on this coverage to keep costs down (Glied, 2000). With the exception of traditional Medicare coverage for non-drug medical services, virtually all insured Americans face managed care policies of some kind. Prior authorization restriction policies are one item in an MCO's toolbox for reducing costs and ensuring appropriate care. In general, MCOs provide insurance coverage for certain medical services and prescription drugs, while excluding others due to cost and/or clinical efficacy. In between these two stances towards coverage is a third option: The MCO will cover the service/drug, but on an *ad hoc* basis, where coverage must be requested explicitly from the insurer *before* the service is rendered or the prescription filled. This contrasts with an MCO's stance towards unrestricted services and drugs, where coverage is implicit in the insurance contract (though sometimes with a modest financial cost to the consumer via a copayment or deductible).

To get permission, the patient's medical provider (rather than the patient herself) must fill out a specific form, provided by the MCO. Authorization forms for prescription drugs generally require the provider to answer some yes-or-no questions regarding why they are choosing to prescribe a restricted drug, particularly when an unrestricted option is available. Forms almost always require documentation of the patient's history of taking drugs in the same class (including the drug itself, if the patient received coverage for it from a previous insurer) to be reported. Generally, the provider will be asked to provide medical documentation of the assertions made in the form. In Appendix B we provide some examples of prior authorization forms used by MCOs. After the form is submitted, the provider and patient must wait until the MCO authorizes it. Authorization requires an administrator at the MCO to review the application and respond accordingly. This generally takes between 1 and 5 business days (American Medical Association, 2017). If the authorization is approved, the patient can then receive the drug or service with coverage. If not, they will not be able to use coverage unless their provider makes another request and receives authorization.

In theory, prior authorization can be required for most medical services and prescription drugs. In practice, since a single hospital stay or physician office visit is comprised of a bundle of many services, requiring prior authorization for some subset of those would be unnecessarily disruptive, forcing providers to deliver care in a piecemeal way. Instead, prior authorization restrictions are generally

applied to discrete services.[5] Prescription drugs, especially specialty and high-cost branded drugs, are the most common treatment to face restrictions, which is why we focus on them in this study. Other commonly-restricted services include surgeries, durable medical equipment, genetic testing, and imaging, most of which are also highly discrete services (America's Health Insurance Providers, 2020). In Section II.D we describe how prior authorization is specifically used in our empirical setting.

The stated purpose of prior authorization restrictions is that they allow an MCO to prospectively review potential drugs whose value, for a specific patient, may not be high enough relative to procurement cost, and reject coverage for prescriptions where this is the case. In this way, coverage under authorization restrictions functions similarly to auto insurance, another setting where insurers require extensive documentation by an expert third party as a prerequisite for coverage. A natural question is, if insurers believe they can observe enough information to make a determination about what care is cost-effective and what care is not, why not make coverage directly contingent on this information in a formulaic way, rather than an ambiguous way requiring human review?[6] First, the implicit formula is likely to be quite complex. For example, for anticoagulants (blood-thinners), newer high-cost non-vitamin K antagonists (so-called "novel anti-coagulants") will be preferred over the low-cost generic blood-thinner warfarin if the patient is already taking some other drug that has an adverse interaction with warfarin (of which a few exist). Codifying all of these requirements might be possible but would be difficult to communicate to non-expert patients, and embedding them into a form for the expert medical provider to fill out may be an easier enforcement mechanism. Second, the patient's provider may be able to observe important signals of the patient's suitability for a specialty drug that cannot be credibly communicated to the MCO, e.g. their ability to tolerate side effects, or interactions with diet and lifestyle. A provider doing paperwork in support of the patient's prescription can thus serve as a costly signal to the MCO of the prescription's merit. In the terminology of Nichols and Zeckhauser (1982), the paperwork serves as an 'ordeal' that screens out prescriptions that providers think are relatively low-value for a given patient.

B. A Model of Prior Authorization Restrictions

We present a simple model of prior authorization restrictions in the typical setting in which they are applied, in the spirit of Finkelstein and Notowidigdo (2019). Consumers face a discrete choice over a high cost drug H, a low cost drug L, and the outside option of purchasing no drug (0). A consumer i gets (social)

[5]For other categories, like inpatient hospitalization, MCOs typically use *retrospective* utilization review instead, rescinding payment for wasteful or fraudulent service provision. Dranove and Spier (2003) present a theory of why MCOs might do this and how it disciplines provider moral hazard. Dunn et al. (2021), Shi (2021), and Eliason et al. (2021) study the effect of review-induced claim denial policies.

[6]i.e., the 'tagging' approach of Akerlof (1978).

value v_i^d from drug $d \in H, L, 0$. Without loss of generality, we normalize $v_i^0 = 0$ for all i such that v_i^H and v_i^L also represent incremental valuation of H and L relative to the outside option. We also define $\Delta v_i = v_i^H - v_i^L$ as the consumer's incremental valuation of H versus L. Similarly, let c^d be the (social) cost of drug d, and again let $c^0 = 0$ and $\Delta c = c^H - c^L$.

The social planner is considering whether to impose a prior authorization requirement on the high cost drug H. The requirement imposes a constant "sludge" cost of a for every consumer whose doctor fills out and submits the authorization paperwork. Not all prior authorization requests are approved, so not every consumer for whom the sludge cost is paid actually consumes the drug. Let θ represent the (exogenously-determined) approval rate.

While the sludge cost imposes a social cost on physicians, insurers, and consumers, it also may affect who ultimately consumes each drug. Let $D_i^H(0)$ represent consumer i's demand for drug H in the absence of the prior authorization requirement, where $D_i^H(0) = 1$ if the consumer gets the drug in the absence of the prior authorization requirement and $D_i^H(0) = 0$ otherwise.[7] Similarly, let $D_i^L(0)$ be consumer i's demand for drug L and $D_i^0(0)$ be consumer i's demand for the outside option. Finally, let $D_i^d(1)$ be consumer i's demand for drug d with the prior authorization requirement in place.

Social welfare without the prior authorization requirement is thus given by

$$SW(0) = \sum_i \left[(v_i^H - c^H) D_i^H(0) + (v_i^L - c^L) D_i^L(0) \right]$$

And social welfare with the prior authorization requirement is given by

$$SW(1) = \sum_i \left[(v_i^H - c^H - \frac{a}{\theta}) D_i^H(1) + (v_i^L - c^L) D_i^L(1) \right]$$

The two key differences between these expressions are (1) a sludge cost is paid for every consumer requesting authorization to fill a prescription for drug H when the prior authorization requirement is in place and (2) different consumers consume the drugs under prior authorization.

To see how these two key differences affect welfare, and to illustrate when welfare will be higher versus lower under prior authorization, we now difference the two expressions. First, we make the following assumptions, which we think of as fairly benign:

1) "No defiers": There are no consumers who consume L or no drug when H is not subject to prior authorization restrictions, but consume H when it is ($D_i^H(0) = 0$ and $D_i^H(1) = 1$)

2) "Independence of irrelevant alternatives": There are no consumers who consume no drug when H is not subject to prior authorization restrictions,

[7] We remain agnostic on the relationship between D_i^H and valuation so as to allow for a variety of behavioral and institutional factors to result in wedges between demand and value.

but consume L when H is subject to restrictions, or vice versal ($D_i^0(0) = 1$ and $D_i^L(1) = 1$, or $D_i^L(0) = 1$ and $D_i^0(1) = 1$)

Given these assumptions, there are five types of consumers. First, there are three groups of inframarginal consumers, those always consuming H, those always consuming L, and those always consuming no drug. Second, there are two groups of marginal consumers: (1) the 'intensive margin marginals' who consume H without prior authorization and consume L when prior authorization is required and (2) the 'extensive margin marginals' who consume H without prior authorization and consume nothing when prior authorization is required. Given this, the difference between $SW(0)$ and $SW(1)$ is given by

$$SW(0) - SW(1) = \underbrace{\frac{a}{\theta}\sum_i D_i^H(1)}_{\substack{\text{Sludge cost for} \\ \text{always H inframarginals}}} + \underbrace{\sum_i (\Delta v_i - \Delta c) D_i^H(0) D_i^L(1)}_{\substack{\text{Incremental social benefit of } H \text{ versus } L \\ \text{for intensive margin marginals}}} + \underbrace{\sum_i (v_i^H - c^H) D_i^H(0) D_i^0(1)}_{\substack{\text{Absolute social benefit of } H \\ \text{for extensive margin marginals}}}$$

The difference in social welfare between the cases without and with prior authorization requirements on H (or the welfare "loss" from prior authorization) consists of three components. The first component is the sludge cost paid for each consumer in the "always H" inframarginal group, amplified slightly by a prior authorization approval rate less than one (θ). This is the pure welfare loss that is the sole focus of most of the literature on administrative costs in healthcare.

The second and third components are the change in welfare due to the shift in social surplus generated by marginals. This is the typical focus of recent papers on administrative burdens in social insurance (Deshpande and Li, 2019; Finkelstein and Notowidigdo, 2019), although our setting contains two margins of substitution rather than one. These components can be either positive (welfare loss due to authorization restrictions) or negative (welfare gain), depending on whether or not incremental or absolute values of drug H for patients on the intensive or extensive margins (respectively) are larger than the incremental or absolute social costs of procuring those drugs.

First, there is a change in welfare due to the shift in the intensive margin marginals from consuming H to consuming L. This component could be posi-

tive (welfare loss due to authorization restrictions) or negative (welfare gain due to authorization restrictions), depending on the incremental social value of the marginal consumption relative to the incremental social cost of that consumption. In cases where H is the branded version and L is the generic version of the same drug, $\Delta v_i = 0$, and this component represents a pure welfare gain to offset the welfare loss from the sludge costs. In other cases, where L and H are imperfect substitutes, the welfare consequences are less clear, as the incremental value could exceed the incremental cost.[8]

Second, there is a change in welfare due to the shift in extensive margin marginals from consuming H to consuming nothing. Again, this component could be positive or negative, but for this group the bar for prior authorization requirements to produce a welfare gain is much higher: Here, the *absolute* social value of the marginal consumption must be less than than the absolute social cost. Again, under full insurance we would expect at least some cases where cost exceeds valuation, especially where the cost of the drug is very high, as is often the case with drugs under prior authorization requirements. But again, we would also expect at least some cases where valuation exceeds cost. Once again, it is an empirical question as to which group will dominate among the intensive margin marginals, an empirical question that we attempt to investigate to the best of our ability below.

When Should Policymakers Restrict Drugs? Our model provides several important insights for where prior authorization is more versus less likely to provide social value. First, prior authorization is more likely to add value in settings where the size of the "always H" inframarginal group is small relative to the size of the marginal groups, since total sludge costs scale with the size of this group. Prior authorization may thus be most useful for "niche" drugs with relatively few users.

Second, prior authorization requirements are likely to be more valuable in settings where the size of the intensive margin marginal group is large relative to the size of the extensive margin marginal group, particularly when a cheaper close substitute exists. As Chandra and Skinner (2012) point out, much of health care spending growth comes from new technologies that have little demonstrated clinical benefit over the status quo. This includes new drugs, and it is thus likely that new, expensive drugs tend to have low incremental value, although they may have large absolute value relative to receiving no treatment.

Rearranging terms, we note that our expression for $SW(0) - SW(1)$ implies that in order for prior authorization to improve welfare it must be the case that:

[8] Under full insurance, we expect there to be some consumers of H for whom incremental cost exceeds incremental value, due to moral hazard (Pauly, 1968). However, the extent to which these consumers make up a large share of drug consumers generally, as well as of the marginals, is an empirical question. In general, we should expect the share of such consumers to be higher in circumstances where the cost difference between L and H is higher.

$$\underbrace{\sum_i \Delta v_i D_i^H(0) D_i^L(0) + \sum_i v_i^H D_i^H(0) D_i^0(1)}_{\text{Social value of all marginal consumption}}$$
$$< \underbrace{\sum_i \Delta c_i D_i^H(0) D_i^L(1) + \sum_i c_i^H D_i^H(0) D_i^0(1)}_{\text{Social cost of all marginal consumption}} - \underbrace{\frac{a}{\theta} \sum_i D_i^H(1)}_{\substack{\text{Sludge cost from} \\ \text{inframarginal} \\ H \text{ consumption}}}$$

This expression realigns our social welfare evaluation into three objects: The drug consumption value effects, drug procurement cost (spending) effects, and sludge cost effects of prior authorization restrictions.

In the following sections of the paper, we estimate the cost of all marginal consumption in the Medicare Part D Low-Income Subsidy Program, as well as the shares of the relevant marginal and inframarginal populations. We then calibrate the per-application sludge costs using cost estimates from the literature. This calibration sets an estimated upper bound on how high the social value of all marginal consumption needs to be in order to make prior authorization as it is currently used in the Part D LIS program welfare-improving. We will then attempt to provide suggestive evidence of the social value of marginal consumption.

II. Setting & Data

A. Medicare Part D and the Low-Income Subsidy

Medicare Part D was introduced in 2006 by the Medicare Modernization Act of 2003 to fill a coverage gap in Medicare, which had previously lacked any outpatient prescription drug benefit. Under Part D, drug coverage is fully outsourced to privately contracted insurers providing coverage on the government's behalf. The Medicare program organizes a centralized market in which beneficiaries may select from one of these private plans. In 2016, Medicare Part D covered about 41 million beneficiaries and accounted for about $94 billion of annual expenditures, of which 86% was paid for by federal and state governments and the remainder by individual beneficiaries through premium payments and cost-sharing.

Part D plans are required to adopt a standardized benefit design, in terms of their financial as well as non-financial features. Specifically, plans follow a standardized cost-sharing structure and are required to cover at least two drugs in each of 148 therapeutic classes. Nonetheless, plans are given some scope for differentiation, in terms of the specific drugs that they cover on their formulary and the specific cost-sharing levels they assign to different formulary tiers. Similarly, plans are given flexibility in setting non-price rationing mechanisms for formulary-covered drugs, such as authorization restrictions, which constitute a focus of this

paper. Meanwhile, consumers enjoy choice among a substantial variety of Part D plans, being free to choose any one of the many Part D plan offered in their service region. In 2016, the average beneficiary had access to 26 stand-alone Part D plans.

While the Part D program is heavily subsidized for all beneficiaries, those with financial need are granted additional subsidies through the low-income subsidy program (LIS), which offers supplemental drug premium and cost-sharing support. Around 30% of Medicare beneficiaries participate in the LIS program. So-called 'dual-eligibles' (those who simultaneously qualify for both Medicare and their state's Medicaid program) are automatically enrolled in the LIS program when they qualify for Medicare, as are beneficiaries in the Medicare Savings Program. Others not automatically eligible for LIS, but who meet income and asset eligibility criteria, can qualify by applying directly.

The generosity of LIS subsidies varies by group, with partial LIS recipients receiving partial premium and cost-sharing support, while the full LIS on which our study focuses receive much more generous subsidies. Specifically, full LIS recipients get a subsidized reduction in their plan premium payments up to the 'benchmark' amount, meaning that those enrolling in a subset of plans (otherwise known as 'benchmark plans') would have no premium responsibility. Multiple benchmark plans are available in a service region; beneficiaries have access to between two and sixteen, with 92% of beneficiaries having at least 5 benchmark plans to choose from (see Appendix Figure A1).

Dual-eligibles in the LIS program additionally receive substantial cost-sharing subsidies: They have all plan deductibles completely waived, under both benchmark and non-benchmark plans. In addition, full LIS beneficiaries in both benchmark and non-benchmark plans are shielded from each plan's drug-specific copayment and coinsurance schedule, which can otherwise be a significant source of out-of-pocket costs for Medicare beneficiaries. Instead, full LIS beneficiaries face their own custom copayment schedule: In 2020, they were charged a copayment of $1.30 for all formulary-covered generic drugs and $3.90 for all formulary-covered branded drugs, though in most cases these nominal copayments are not actually collected. In addition to substantially reducing out-of-pocket costs, this policy also makes plans effectively uniform in their financial characteristics for dual-eligible beneficiaries, nullifying any variation in cost-sharing determined by the plan sponsor.

Given that the full-LIS population's out-of-pocket expenses are uniform for covered drugs, plans most materially differ in the set of drugs covered by their formularies (the set of drugs that they offer coverage for), along with the utilization management requirements imposed on those drugs that are covered. As a result, this setting is particularly well-suited for isolating the impact of prior authorization and formulary restrictions on different drugs. Note that in this context, formulary exclusion of a drug means a beneficiary would have to pay the full sticker price of that drug out-of-pocket if they opt to purchase the drug. Because

Part D regulations require plans to cover at least two drugs in each therapeutic class, beneficiaries will generally have another covered drug that they can switch to as a substitute, although it might be an imperfect substitute given variation in how good

This heterogeneity is important, as it leads to substantial cross-plan variation in the set of drugs covered by plans' formularies, as well as the prior authorization restrictions imposed on covered drugs, all of which we leverage as part of our study. To take the popular anti-cholesterol drug Lipitor as an example, of the nine benchmark plans in New York in 2009, six plans covered the drug on their formulary while three did not. Further, even among the six plans that did cover the drug, two required prior authorization for beneficiaries to obtain coverage, while four did not. As such, if a LIS beneficiary in New York wanted to fill a prescription for Lipitor but was on one of the three plans that did not cover it, they had the option to either pay for its cost 100% out-of-pocket, substitute to another drug, or switch plans entirely. Given that different anti-cholesterol medications (as well as many other classes of important drugs) are not perfect substitutes for one another, these cross-plan differences in drug coverage can be meaningful for beneficiaries and carry real consequences.

B. Data

We make use of several administrative datasets from the Centers for Medicare and Medicaid Services (CMS). These data contain information on beneficiary program enrollment status, medical utilization, and prescription drug utilization within the Medicare program. The data is nationwide in scope and extends from 2007 to 2015, tracking drug utilization for all Medicare beneficiaries and medical utilization for all beneficiaries outside of Medicare Advantage.

Beneficiary Demographics, Enrollment, and Choice Status.

We obtain information on beneficiary demographic characteristics and plan as well as program enrollment from the Medicare Beneficiary Summary File. This file provide important demographic information such as age, gender, and geographic location, including the Part D plan region to which individuals belonged that year. It additionally tracks enrollment status at a person-month level for different Medicare programs, including Part A (the hospital benefit), Part B (outpatient), Part C (Medicare Advantage), and Part D (prescription drugs). These Medicare files also track enrollment in the LIS program at a person-month level, and whether beneficiaries qualify for the full LIS subsidy.

We combine this data with a newly released plan election type file. This file covers all Part D enrollment spells from 2007-2015, and for each spell tracks whether enrollment was initiated through active choice or default auto-assignment. Importantly, in addition to listing the plan a beneficiary was enrolled in during each month, the file also includes the default plan that was assigned to the beneficiary,

even if the beneficiary opted out of that default. This allows us to observe the *assigned* plan as well as the *realized* plan for each beneficiary.

Plan Characteristics and Formulary Data.

We obtain information on plan characteristics from publicly available CMS datasets, which cover all Part D plans offered during our sample period. These data track the set of plans offered in each Part D plan region. For each plan in each year it was offered, we are able to observe the monthly premium that the plan charged (both the Medicare-paid portion and the beneficiary-paid portion) and the plan's benchmark status.

We additionally obtain drug-level formulary data for each Part D plan, publicly available from CMS, information which is key to our study design. This data tracks the set of drugs covered by each plan's formulary, at a drug-by-drug. For each covered drug, the data also indicates the type of utilization restrictions imposed by the plan on the covered drug, including prior authorization, step therapy, or quantity limits.

While the original CMS data defines drugs at an NDC level, we revise the drug definition to instead be at the combination of active ingredient (e.g., atorvastatin, or warfarin) and brand/generic status, and then aggregate the data up to this level. In doing so, we effectively treat different doses and different modes of administration as equivalent; if a formulary covers at least one dose or mode of administration, we count every possible dose/mode of administration as covered. Similarly, this approach also means that we treat identical generic substitutes as equivalent, and treat the full set of generic substitutes as covered so long as at least one is by a plan. We do this because we had noticed that the formulary data, if taken at face value without these adjustments, would have implied that many drugs that we observed patients frequently using (and which their insurer paid for) were excluded from coverage.

Outpatient Prescription Drug Data.

We track outpatient prescription drug usage for a random 20% sample of Part D enrollees using claims-level Medicare data from the Part D Event files. Each claim represents an event where a beneficiary filled a single prescription of a given drug. For each claim, we observe the specific drug prescribed and filled (at the NDC code level), the quantity/days supply for the fill, as well as the date when the prescription was filled, and the cost charged to the beneficiary and to the Medicare program. As with the formulary data, rather than defining unique drugs directly based on NDC codes, we instead define a unique drug based on the combination of active ingredient and brand/generic status. Setting comparable drug definitions across these two data files ensures that we can cleanly link the formulary file to the outpatient drug claims at an individual drug level, to then examine the impact of formulary coverage/prior auth status on a given drug's utilization.

For our main analyses, we restrict only to drugs that were listed as covered by at least one Medicare Part D plan formulary in that calendar year. This is meant to remove uncovered drugs from our sample, for which there would be no coverage variation, and additionally to remove miscellaneous drug types whose coverage status we would not be able to track in formularies whatsoever.[9] As part of our analyses, we aggregate utilization statistics from the drug claims data up to the plan-drug-market-year level, for our main analytic sample of interest. Plan for these purposes is defined based on beneficiaries' originally assigned, rather than actual enrolled, plan in the listed contract year. We then divide the aggregated statistics up by the number of people in a given plan-drug-market-year cell in our analytic sample, to yield per capita utilization statistics. Aggregating this data at a plan-level rather than individual beneficiary level substantially reduces the number of observations in the data, and in doing so makes analysis of the data much more tractable.

C. Sample Selection

For our main analyses, we employ a single subsample of LIS beneficiaries, which is defined to cleanly and maximally exploit natural experiments in exposure to prior authorization. To start, we restrict to those in Medicare Parts A, B, and D, and not enrolled in Medicare Advantage. We further restrict to individuals who qualify for the full rather than partial LIS subsidy, who would qualify for the full premium subsidy and for whom cost-sharing for formulary-covered drugs would be minimal. We generally sample at the beneficiary-year level and require these restrictions to be true for every month in a year in which we include a beneficiary in our sample.

Critically, we additionally restrict to individuals who were randomly re-assigned from their incumbent plan to a new plan; these individuals are made up of previous non-active choosers, whose incumbent plans lost benchmark status. In doing so, we exclude incumbent plan-market combinations where reassignment is expected to be non-random, based on reassignment rules. For example, reassignment will not be randomized if the carrier of the incumbent plan offers another benchmark plan in the market, as all reassignees will simply get funneled to that other plan.[10] Finally, for beneficiaries whose assigned plan retained benchmark status for two full years following the beneficiary's reassignment, we include data for two years post-assignment. For beneficiaries whose assigned plan lost benchmark status in the second year post-assignment, we drop the second year and only keep observations from the first year. Table 1 shows summary statistics for the plans included in our sample. There are 1,444 plan-years in our sample, with an average of 804

[9] For example, our formulary dataset generally does not track coverage status for over-the-counter (OTC) drugs.

[10] We perform additional robustness checks to validate that reassignment out of different incumbent plan-market combinations is randomized, and drop combos from the sample from which reassignment does not appear random, based on observables.

beneficiaries per plan. The average plan requires prior authorization for 12% of drugs it covers.

Table 1—: Included plan summary statistics

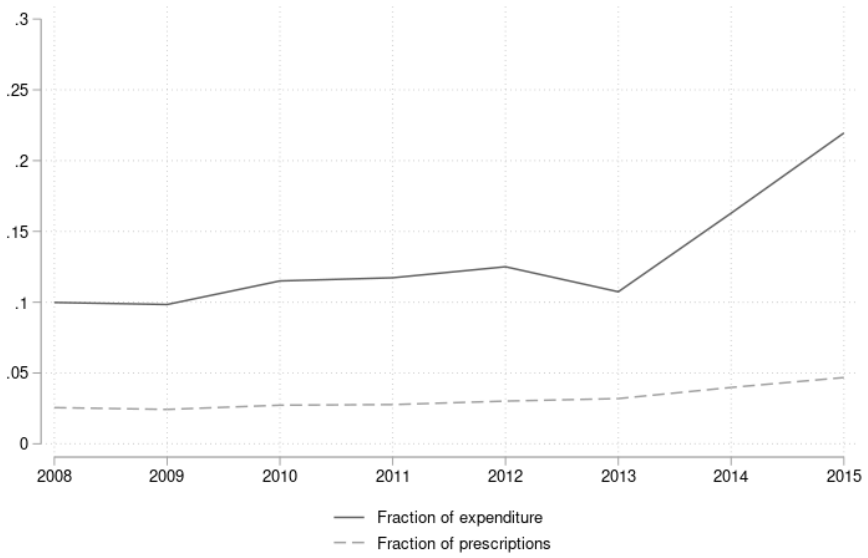
	Whole sample	First year	Second year
Plan-years	1,444	1,071	723
Mean benes per plan	804.1	655.2	635.3
Mean % of drugs under prior auth	12.0	11.7	12.6
	(4.5)	(4.7)	(4.0)
Mean % of drugs excluded	28.4	27.4	29.6
	(9.8)	(10.6)	(8.3)

Note: Observations are at the plan-year level. The first column includes all beneficiary-years in our sample. The second column includes all beneficiaries in the first year after being re-randomized to a new plan. The third column includes beneficiaries whose randomized plan retained benchmark status for two full years following initial reassignment.

D. Prior Authorization in Medicare Part D

Before proceeding to our main empirical analysis, we describe the use of prior authorization restrictions in Medicare Part D over time and across drug types. Figure 1 shows the use of prior authorization restrictions for a 20% sample of Medicare Part D claims from 2008-2015. Use of prior authorization increased over this period, and by 2015 5% of all claims under part D involved a prior authorization requirement, accounting for more than 20% of overall spending. Use of prior authorization differs substantially by therapeutic class. Appendix Table A1 shows the frequency of prior authorization restrictions for the top 30 therapeutic classes by Part D drug expenditure in 2008-2015. These classes together make up 83% of all spending. Among the highest spending classes, prior authorization is particularly common for biological response modifiers (affecting 70% of total claims spending), immunosuppressants (66%), and anti-neoplastic drugs (58%). Prior authorization is also frequently applied in mental health treatments, affecting 7% of anti-psychotic spending and 8% of antidepressant spending. On the other hand, prior authorization is less common for important classes like the antihyperlipidemic drugs (including well-known 'blockbuster drugs' like Lipitor and Crestor) and insulins. Table 2 shows use of prior authorization restrictions for all drugs in our sample. The average drug is under prior authorization restriction for 13% of plan-years. Branded drugs without generic equivalents are under prior authorization restrictions for 24% of plan-years, much higher than either generics or branded drugs with generic equivalents. Branded drugs with generic substitutes are on average excluded by more than half of plan-years, and exclusion is also common for branded drugs without generics. The price of branded

Figure 1. : Use of prior authorization in Medicare Part D



drugs[11] without generic equivalents is higher than for branded drugs with generics or generics. Figure 2 shows that higher cost drugs are more likely to have prior authorization restrictions. This pattern is observed for both branded and generic drugs, but is more pronounced for branded drugs, with 60% of drug-years in the top ventile of price under prior authorization.

III. The Effect of Authorization Restrictions on Drug Utilization

We begin by estimating the effect of prior authorization restrictions on drug utilization at the person-drug level. We specify that as estimating the treatment effect of moving a drug from having coverage with no restrictions to having coverage with restrictions on the utilization of that same drug, all else equal. Our initial analysis seeks to estimate the average treatment effect, averaged over beneficiary-drug pairs, $E_{id}[Y_{id}(1) - Y_{id}(0)]$.

A. Research Design

Before delving into our research design, it is worth highlighting what the ideal experiment would look like. Ideally, beneficiaries would be randomly assigned to a formulary where any given drug could either be covered with restrictions

[11] 'Price' is defined as claim total cost divided by days supply of the drug. This ignores any rebates the insurer may receive from drug manufacturers.

Table 2—: Formulary restrictions by drug type

	All	% Drug type		
		Branded without generic	Branded with generic	Generic
Number of drug-years	12,727	4,457	3,443	4,736
Number of unique drugs	2,017	847	609	742
% of plan-years under prior auth	12.6	23.5	5.9	7.8
	(23.9)	(30.4)	(15.0)	(18.6)
% of plan-years excluded	29.8	27.5	57.4	10.2
	(34.9)	(30.7)	(37.7)	(18.2)
Price per day supply (USD)	26.0	61.2	16.1	5.6
	(128.4)	(218.6)	(44.1)	(22.3)
% of benes with any use	0.8	0.3	0.2	1.7
	(2.9)	(1.1)	(1.0)	(4.4)
Cost per enrolled bene (USD)	2.6	5.0	1.3	1.4
	(10.3)	(16.0)	(5.9)	(3.5)

Note: A 'drug' is defined as a combination of active-ingredient and whether the product is branded/generic. Products containing different doses of the same active ingredient and with different modes of administration are all counted as the same drug.

or without restrictions, with all possible formulary arrangements available for assignment. We would then be able to estimate the simple regression of utilization on restriction status,

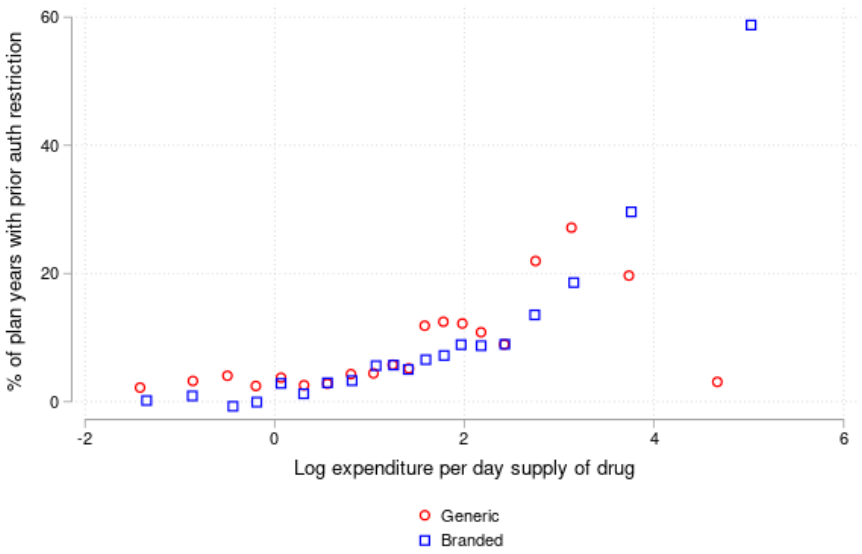
$$Y_{id} = \beta \text{Auth}_{id} + \epsilon_{id}$$

and β would represent an unbiased estimator of the average treatment effect. Unfortunately, this is not a proper representation of our setting. Our setting has random assignment to default auto-enrollment into one of a handful of insurance plan options. This creates five complications.

First, drugs can be unrestricted and restricted as well as fully excluded by the plan, a third category. Because we are interested in the effects of prior authorization restrictions relative to the counterfactual where the drug is included on the formulary and not restricted, comparing restricted drugs to all other drugs would not provide an estimate of our parameter of interest because it would improperly conflate unrestricted and excluded drugs. There are thus two possible treatments, restriction and exclusion. We choose an approach where we control for whether a drug is excluded, focusing on estimating the effect of restriction relative to unrestricted coverage.

Second, not every formulary is available for assignment. Because there are 3 potential treatments and approximately 2000 unique drugs, this would require

Figure 2. : Prior authorization restrictions by drug price



$3^{2000} = 1.7 \times 10^{954}$ unique formularies to be available, which is larger than our sample size. Moreover, as shown in Section II.D, drugs facing authorization restrictions are non-randomly chosen, and tend to be expensive on-patent branded drugs. Therefore we need to control for drug fixed effects, or else our treatment effect will be biased by differences in counterfactually-unrestricted levels of use among frequently-restricted drugs.

Third, the random assignment we use is not to *formularies* but instead to *plans*. As we describe in Section II.A, for LIS beneficiaries, plans have limited scope to differentiate on dimensions other than their formulary. However, we need not assume this is true. Instead, we absorb differences across plans that secularly affect drug utilization, using plan fixed effects.[12]

Fourth, the randomization we use is across plans within a specific market (defined by Part D region and year), with the formularies available varying across markets. If this variation reflects the tastes of the market (e.g., drugs are less likely to be restricted by plans in a given market if beneficiaries prefer them in that market), this will introduce correlation between treatment assignment and unobservable preferences for a drug. We therefore perform all of our analysis within-market, interacting all fixed effects with fixed effects for the market. That is, instead of controlling for drug and plan fixed effects, we control for drug-market

[12]In later analyses, identification requires us to drop these plan fixed effects. In Appendix Table A6, we show that our results are robust to excluding them, implying that drug exclusion and restriction fully capture the effect of plans on drug consumption.

and plan-market pair fixed effects.

Fifth, we need to account for spillovers across drugs. Since some drugs are substitutes, restrictions on one drug may have positive effects on the consumption of another. While this fact alone would not lead to bias when estimating the direct effect of prior authorization on usage of the focal drug, it will if prior authorization restriction status is correlated across substitute drugs. For example, take a situation in which three drugs can be used to treat a condition: Drug A, Drug B, and Drug C (ordered by increasing price). Generally, restriction status is reserved for the most expensive drugs. If an MCO restricts Drug B, they are likely to also restrict (or exclude) Drug C. Therefore, if we didn't control for the status of Drug C, we would underestimate the effect of prior authorization applied to B alone on consumption of B, which would be confounded by the (positive) effect of restricting access to C. In other words, we want to compare the use of a drug for beneficiaries assigned to a formulary where that drug is under prior authorization restriction versus beneficiaries assigned to a formulary where that drugs has no restriction, *holding fixed the formulary status for all substitute drugs.* Since there are roughly 2000 drugs, controlling richly for all drugs' substitution effects on others would require us to estimate approximately 4 million parameters, which is infeasible given our limited number of unique formularies. Instead, we construct a single control for prior authorization status of substitute drugs, $\text{Auth}_{id}^{Sub} = \sum_{k \in \mathcal{C}(d)} s_k \text{Auth}_{ik}$, i.e., the share of drugs k within the same class $\mathcal{C}(d)$ that face authorization restrictions, weighted by their market shares s_k in the entire sample.[13] We also control for a similar measure for formulary exclusion of substitute drugs.

Incorporating these issues, our primary estimating equation is

$$
\begin{aligned}
(1) \qquad Y_{id} &= \beta \text{Auth}_{id}^{\text{Assigned}} + \delta \text{Excl}_{id}^{\text{Assigned}} + \alpha_{dm(i)} + \eta_{p(i)m(i)} \\
&+ \beta^{Sub} \text{Auth}_{id}^{\text{Sub,Assigned}} + \delta^{Sub} \text{Excl}_{id}^{\text{Sub,Assigned}} \\
&+ \alpha_{dm(i)} + \eta_{p(i)m(i)} + \epsilon_{id}
\end{aligned}
$$

Our approach, which relies on two-way fixed effects, is similar in spirit to difference-in-differences designs used in the literature (de Chaisemartin and D'Haultfœuille, 2020). β is identified off of many two-drug, two-plan comparisons, comparing utilization differences across plans for a drug whose formulary status varies across the two plans to one whose formulary status does not vary.

Our approach is valid under a handful of assumptions. First, since we are leveraging the auto-assignment mechanism, it must be that assignment to authorization restrictions is truly random (within a market), i.e., for a given drug d, the formulary status of that drug in the plan beneficiary i is assigned to must be in-

[13] This weighting scheme is exactly appropriate if demand is logit or nested logit, under which $\frac{\partial Y_{id}}{\partial \text{Auth}_{ik}} \propto s_{ik}$

dependent of their unobserved propensity to use the drug, ϵ_{id}. In the next section we provide a number of balance tests which support this assumption. Second, we require an approach akin to the 'parallel trends' assumption in difference-in-differences research designs: Since we identify our treatment effect using two-way drug and plan fixed effects, it must be that plans that restrict a drug do not engage in other, unobservable actions that encourage or dissuade beneficiaries from using that drug (relative to plans that do not restrict the same drug). Plans can have secular effects on drugs overall (e.g. having restrictive pharmacy networks that make it difficult to pick up drugs), which will be absorbed by the plan fixed effects, but are assumed to not *differentially* affect drugs (e.g. by instituting other drug-specific coverage policies) in ways that are correlated with use of prior authorization restrictions. For LIS beneficiaries, plans have limited capacity to affect drug utilization beyond formulary design, so we feel comfortable making this assumption.

B. First Stage

Before we estimate our primary regression, we first estimate a 'first-stage' regression to show that compliance with the assigned default is extremely high. We perform basic checks of compliance and balance. For more detail on when and why beneficiaries comply with default assignments in this setting, see Brot-Goldberg et al. (2021).

We estimate the following regression to determine compliance with the assigned default:

$$
\begin{aligned}
(2) \qquad \left\{ \begin{array}{c} \text{Auth}_{id}^{Enrolled} \\ \text{Excl}_{id}^{Enrolled} \end{array} \right\} &= \vec{\gamma_1}\text{Auth}_{id}^{Assigned} + \vec{\gamma_2}\text{Excl}_{id}^{Assigned} \\
&+ \vec{\gamma_3}\text{Auth}_{id}^{Sub,Assigned} + \vec{\gamma_4}\text{Excl}_{id}^{Sub,Assigned} \\
&+ \alpha_{dm(i)} + \eta_{p(i)m(i)} + + u_{id}
\end{aligned}
$$

where the (vector) coefficients $\vec{\gamma_1}, \vec{\gamma_2}$ measure compliance with assignment at the drug level. In practice, the input dataset into this regression as well as our main regression contain approximately 1.6 billion observations, since each observation is an individual-drug-year tuple. Estimating such a model with high-dimensional fixed effects is computationally burdensome. To ease computation, we collapse our dataset down to the drug-plan-market (the assigned treatment unit) level and measure averages of outcomes over these aggregates. This is equivalent to clustering our standard errors at the drug-plan-market level. Indeed, we eventually cluster at the plan-market level, so we do not lose any information through this shortcut. We weight plan-market observations by the number of assigned beneficiaries.

Estimates from Equation 2 are given in Table 3. Our first stage is extremely strong, with F-statistics of over 67 million, well above the usual threshold of

10. Approximately 97% of beneficiaries comply with their default assignments. One worry is that, although this result is strong for the general population, most individuals do not use most drugs. If compliance is primarily centered around beneficiary-drug pairs where the beneficiary is unlikely to use the focal drug, our instrument will only be strong for an irrelevant population. To try to address this, we re-estimate the first stage on a subset of beneficiary-drug pairs where the beneficiary took the drug at least once in the prior year. These beneficiaries should be especially likely to take the drug again in the following year. We present results from these regressions in Appendix Table A3, which line up well with those in Table 3. This is unsurprising given that Brot-Goldberg et al. (2021) show that beneficiaries in this setting do not actively choose plans in response to defaults that exclude their previously-used drugs.

Table 3—: First Stage Regressions

	$\text{Auth}^{\text{Enrolled}}$	$\text{Excluded}^{\text{Enrolled}}$
$\text{Auth}^{\text{Assigned}}$	0.969	< 0.001
	(< 0.001)	(< 0.001)
$\text{Excluded}^{\text{Assigned}}$	< 0.001	0.969
	(< 0.001)	(< 0.001)
F-statistic	71,816,813	67,609,323
Number of market-years	210	
Number of drug × beneficiary-years	1,613,559,179	
Average plans per market-year	6.6	
Average beneficiaries per plan-year	4,956	

Note: Observations are at level of drug × assigned plan × enrolled plan × year. Column (1) shows estimates from regressions of prior authorization of drug d in enrolled plan on prior authorization and exclusion status of drug d in assigned plan.

Another worry is that the default assignment we leverage is not truly random. We perform three sets of balance tests. First, we measure a placebo 'first stage', replacing the dependent variables in Equation 2 with their equivalents from the *prior year*, which should not be affected by contemporaneous restrictions. Second, we measure a placebo outcome regression, estimating the 'effect' of prior authorization on utilization in the prior year, using the specification in Equation 1. Third, we measure the 'effect' of prior authorization on beneficiary characteristics (gender, race, and Elixhauser comorbidity index). We display the results from these exercises in Appendix Table A4. Reassuringly, our estimated differences from these regressions are substantively small, though often statistically significant.

C. Main Estimates

We now estimate the effect of prior authorization restrictions on utilization by estimating the regression in Equation 1. Since our 'first stage' coefficient is so close to 1, we eschew an instrumental variables approach and simply estimate the intent-to-treat effect. We focus on two outcomes: Total spending by that beneficiary on the drug in that year, and a binary indicator for whether the beneficiary filled the drug at least once in that year. We estimate two versions of these regressions: One without the controls for restrictions on substitutes (in the first and third columns), and one with those controls (in the second and fourth columns). For all regressions, we continue to cluster standard errors at the plan-market level. We present these results in Table 4.

Table 4—: Main Regressions

	Spending		% Ever filled	
$\text{Auth}^{\text{Assigned}}$	-0.748	-0.767	-0.091	-0.098
	(0.036)	(0.037)	(0.003)	(0.003)
$\text{Excluded}^{\text{Assigned}}$	-0.951	-0.968	-0.119	-0.123
	(0.0320)	(0.032)	(0.0031)	(0.0031)
Auth^{sub}		0.155		0.049
		(0.029)		(0.004)
$\text{Excluded}^{\text{sub}}$		0.519		0.139
		(0.041)		(0.008)
F-statistic	514	269	781	420
R^2	0.831	0.831	0.978	0.978
Control Mean	3.555		1.307	
Reweighted Control Mean	3.613		0.403	
Number of drug × plan years	2,141,993			
Number of market years	210			
Number of drug × bene years	1,732,564,415			
Average plans per market-yr	6.6			
Average benes per plan	807			

Note: This table presents estimates from regressions of utilization measures on prior authorization and exclusion of a given drug and other drugs within the class. Each observation is a drug-plan-year. Regressions include plan-market-year and drug-market-year fixed effects. Prior authorization of substitute drugs is mean prior authorization status of all other drugs within the class, where drugs are weighted by their average expenditure across all plans in the sample.

While our results seem small in magnitude, this is because the overwhelming majority of beneficiary-drug-year tuples result in zero utilization. Therefore, we need to reference a control mean to properly situate the results. We construct a control mean by estimating a regression of $Y_{idt} = \alpha_{dm(it)} + \eta_{p(it)m(it)} + \epsilon_{idt}$ only on control observations. We then average over the α_{dm} values to construct an ap-

propriate mean spending/utilization level estimate for untreated beneficiaries and list this as "Control Mean" in Table 4. However, this is still the wrong reference mean, since our treatment effect is, implicitly, a weighted average treatment effect where weights are proportional to treatment variance in assigned plans. Therefore, we reweight the α_{dm} values by $\text{Var}[\text{Auth}_{id}|d, m]$, the variance of authorization restriction status for a given drug in a given market across beneficiaries. Since drugs with lower baseline utilization are weighted more heavily in our regression, we must account for that appropriately.

Given this re-weighted control mean, our results indicate that prior authorization restrictions reduce spending on the focal drug by 21.2%, and reduce the share of beneficiaries who ever use that drug by 23.6%.[14] Restrictions on substitute drugs, as expected, raise utilization of the focal drug by an economically meaningful amount. Despite the high unconditional correlation between restrictions on the focal drug and its substitutes, including these controls has a minimal effect on our estimated coefficients, due to us conditioning on drug and plan fixed effects. This result is robust to other utilization measures. In Appendix Table A5 we estimate the same regression using two other measures of utilization: The number of fills during the year (reduced by 25%), and the total days supply for the year (reduced by 28%).

D. Heterogeneous Effects

The point of prior authorization is to deter low-value care. We therefore care about which care is being deterred, and for whom. We get at this by estimating the effects of prior authorization on utilization across various drug and beneficiary categories.

We begin by examining heterogeneous responses by patient characteristics. Recent work has found disparities in health care treatment by gender Cabral and Dillender (2022) and race [CITE?]. Prior authorization requires physicians to exert effort on behalf of their patients; if they are generally less-willing to exert effort on behalf of minority patients, prior authorization may inefficiently deter care for those patients. We separately estimate effects for female vs. male beneficiaries, as well as for white vs. non-white beneficiaries. We report these results in Table [X]. While we find large level gaps in the effect, with women facing a larger deterrence effect, when we rescale the level effect to account for the fact that women tend to consumer more drugs generally, we see roughly-equal effects in percent terms. However, we see that non-white patients experience a greater deterrence effect even in relative terms.

We do the same exercise, segmenting beneficiaries by their health status, measured by quartiles of the Elixhauser Comorbidity Index, a count of the total number of chronic conditions the beneficiary has, measured in the year prior to

[14]The "% ever filled" indicator is the number of beneficiaries per 100 who ever fill a prescription in a year for the focal drug. The reweighted control mean of 0.403 indicates that the (weighted) average drug is ever used by 0.403% of the population.

the index year from medical claims data. We present these results in Table [X]. We see that sicker beneficiaries experience larger deterrence effects of prior authorization, both in levels and proportionally. This, paired with the above results, is worrisome: It suggests that prior authorization is targeting characteristics which we would expect it not to intend to target.

We also study the effects by different drug categories, which we display in Table [X]. We see that effects are roughly similar in relative terms across chronically-used and acutely-used drugs, albeit with slightly larger effects for the former. We also estimate effects for a subset of drugs evaluated by the National Institute for Health and Care Effectiveness (NICE), an organization in the United Kingdom that evaluates prescription drugs on their cost-effectiveness to determine regulation under their National Health Service. NICE has three categories: 'Recommended,' meaning that NICE generally recommends use of the drug for its intended purpose; 'Limited recommendation,' meaning that NICE only recommends the drug for certain patients; and 'Not recommended,' meaning that NICE does not recommend that physicians ever prescribe the drug. (There is a fourth category: Those drugs which NICE has not evaluated) We split drugs across these four groups and estimate group-specific average treatment effects. We find substantially larger relative effects of prior authorization on 'not recommended' drugs compared to 'recommended' or 'limited recommendation,' consistent with the idea that prior authorization is being used to exclude cost-ineffective uses of expensive drugs.

IV. Substitution Patterns and Spending Effects

While our above results show that prior authorization reduces use of restricted drugs, this is only partially sufficient to understand its effects on overall spending. To estimate these effects, we need to estimate the extent to which patients who are deterred from using a restricted drug substitute to an unrestricted drug or to no drug at all.

Determining the importance of substitution versus extensive margin effects in a reduced-form way would be challenging: It would require us to estimate the effect of prior authorization for a given drug on every other drug. As mentioned in Section III.A, our setting has too many drug options and not enough unique formularies for this to be a viable option. Therefore we must enforce some structural restrictions on drug demand to properly estimate substitution patterns and perform this decomposition.

We assume that beneficiaries and their prescribing medical providers choose a drug by maximizing a utility function representing their joint decision-making process (c.f. Ho and Pakes (2014) and Brot-Goldberg and de Vaan (2018) for similar approaches to modeling joint decision-making). We follow recent work by Dubois et al. (2019) and Ganapati and McKibbin (2019) and assume that the pair chooses a single drug (or no drug) to consume each year from a therapeutic class. That is, by this assumption, we assume that drugs in *different* therapeutic

classes are neither substitutes nor complements and we assume that drugs in the *same* therapeutic class are potentially substitutes *but are never complements*. We assume that the utility function to maximize is:

$$u_{idt} = \underbrace{\beta_{c(d)}\text{Auth}_{idt}^{Assigned} + \delta_{C(d)}\text{Excl}_{idt}^{Assigned} + \alpha_{dm(it)}}_{V_{idt}} + \epsilon_{idt}$$

We allow beneficiary-provider pairs to have preferences that vary in mean terms across drugs, and allow pairs in different markets to have different preferences across drugs ($\alpha_{dm(i)}$). Pairs face a constant barrier (β and δ, which vary across therapeutic classes C) to getting a drug depending on its formulary status. We normalize u_{i0}, the mean utility of the outside option of getting no drug, to zero.

Finally, we assume that unobserved preferences for options ϵ are drawn from a standard Gumbel distribution, independent and identically distributed across beneficiaries, but correlated across drug options within a beneficiary, with correlation $1 - \lambda$. This structure gives rise to a nested logit demand system, with a nest for all 'inside goods' (the choice to take a drug within the class rather than take no drug). That is, the probability of taking a drug d is

$$P_{id} = \frac{\exp \frac{V_{idt}}{\lambda} \left(\sum_{k \in C} \exp \frac{V_{ikt}}{\lambda} \right)^{\lambda - 1}}{1 + \left(\sum_{k \in C} \exp \frac{V_{ikt}}{\lambda} \right)^{\lambda}}$$

and the probability of taking no drug is

$$P_{i0} = \frac{1}{1 + \left(\sum_{k \in C} \exp \frac{V_{ikt}}{\lambda} \right)^{\lambda}}$$

This nesting structure is essential, because the standard conditional logit substitution effects depend on the share of relevant beneficiaries taking each option. Since most beneficiaries take no drug, this would lead us to incorrectly predict substantial extensive-margin substitution in response to authorization restrictions. Instead, the extent to which beneficiaries substitute on the intensive margin (to another drug) or the extensive margin (to no drug) depends on $\lambda \in [0, 1]$, to be estimated. Lower (higher) λ values imply that relatively more substitution is on the intensive (extensive) margin.

We note that, in contrast to our earlier reduced-form regressions, our drug choice model omits plan fixed effects. This is due to identification constraints in the nested logit. As Berry (1994) and Berry and Haile (2016) note, in the nested logit, an econometrician needs two instruments for each product: One for the price (in this case the presence of prior authorization restrictions), and one for the quantity of use across the nest (i.e., the total share of beneficiaries who use *any* drug). Our implicit instrument for a drug's own authorization restriction for a beneficiary is whether the beneficiary was assigned to a plan with a restriction; our

instrument for the group share is the full regime of prior authorization restrictions facing the beneficiary, including both the drug in question and all other drugs. In a simplified setting with two drugs, H and L, where only H ever faces prior authorization restrictions, both the coefficient on prior authorization and the nesting parameter are identified from changes in outcomes in response to H facing restrictions; the coefficient on prior authorization is determined from changes in H's market share; whereas the nesting parameter is determined from changes in L's market share. Since a plan is defined by its formulary, we would only be able to identify differences across plans *within a formulary*, which is arbitrarily defined across classes. We instead simply omit these fixed effects. In Appendix Table A6 we show that our reduced-form results are the same without plan fixed effects, alleviating any worries about omitted variable bias.[15]

To construct the relevant data set, we define a beneficiary as taking a drug within a class if they ever filled it during the year. For beneficiaries who took multiple drugs within a class, we assign them to the drug they filled with the highest days supply during the year, and break ties randomly. We restrict to classes where drugs under prior authorization made up at least 3% of spending, and where the average spending on the drug per beneficiary-year in our data exceeded $1. This left approximately 60 therapeutic classes. We estimate the model using maximum likelihood estimation, and do so separately for each therapeutic class.

Within each therapeutic class, we restrict only to region-year pairs where we observe that (1) at least two drugs were ever taken, and (2) at least one faced variable prior authorization status (i.e., was restricted in at least one plan and unrestricted in at least one plan). Restriction (2) is required for us to identify direct effects of prior authorization, whereas (1) is required for the identification of λ.

Our parameterization requires us to estimate over 300,000 fixed effects, which is computationally burdensome in typical optimization packages. Moreover, our setting has substantial sparsity in market shares, in that there are many drug-plan pairs for which no beneficiaries in the plan are taking the drug in a given year; this makes the inversion approach of Berry (1994) impossible. Instead we exploit the equivalence between conditional logit estimation and Poisson regression (Guimarães et al., 2003) and recent computation innovations in Poisson regression with high-dimensional fixed effects (Correia et al., 2020) to estimate the model.[16]

[15]The irrelevance of plan fixed effects is not too surprising here, given that plans have limited scope to differ in ways not captured here already by their use of prior authorization and exclusion.

[16]Specifically: We collapse our data down to utilization counts at the plan-drug-region-year level. We can then recover conditional logit estimates by estimating a Poisson regression model of the counts on the arguments in the drug-specific utility function, as well as a fixed effect for the plan-class-region-year, which absorbs the differences across plans in their denominators. The methods cited are for the conditional logit rather than the nested logit. To implement the nested logit, we transform the problem into two conditional logit estimation problems. First, we estimate $P_{id|d\neq 0}$, which is a conditional logit expression. Then, we estimate $P_{i(d\neq 0)}$ by regressing the counts of beneficiaries in the plan-region-year who took any drug in the class on the plan-class-region-year-specific inclusive value of taking any drug, with the coefficient on that regression identifying λ.

We compute standard errors using a bootstrap procedure.

With the model estimated, we can then use it to ask how the current use of authorization restrictions affects total spending and utilization. Using our model, we simulate demand for drugs under 1) the status quo assigned formularies; and 2) an alternative arrangement where we remove all uses of authorization restrictions (but otherwise leave formulary exclusion as it is).[17] We then compute the differences between these two simulations and display them in Table 5. We compute spending by assuming that beneficiaries who consume a drug spend an amount equal to the empirical average amount spent on that drug by beneficiaries who consumed it in the same region-year pair.[18]

Our results suggest that prior authorization reduced drug spending by 3.1%, approximately $61 per beneficiary-year. This spending reduction is composed of a 24.8% reduction in utilization of restricted drugs ($74 reduction in spending per beneficiary-year), and an 0.7% increase in the utilization of unrestricted drugs ($13 increase in spending per beneficiary-year). Our current estimates show that just under half of beneficiaries switch to another drug rather than substitute away from drugs altogether. This is likely to be undesirable, unless the full value of the drug is below its full cost (in contrast to whether the incremental benefit is above the incremental cost).

Table 5—: Counterfactual Simulations

	Total	Restricted Drugs	Unrestricted Drugs	No Drug
Change in	-3.1%	-19.4%	+0.7%	-
Spending	-$61	-$74	+$13	-
Per Capita				
Change in	-0.9%	-24.8%	+0.7%	+0.1%
# Users	-0.04	-0.08	+0.04	+0.04
Per Capita				
Diversion	-	-100%	44.5%	55.5%

Note: This tables presents results from an exercise where we switch beneficiaries from facing no authorization restrictions to facing the status quo. The first two panels detail the change in spending and utilization of all drug, restricted drugs (those drug-plan-region-year observations where an authorization restriction was in place in the status quo), unrestricted drugs, and no drug. In those panels, the upper row gives the percent change in these quantities, while the lower row presents the absolute change per beneficiary-year. The final panel details the share of beneficiaries moving away from restricted drugs to either unrestricted drugs or no drug.

[17] Note that beneficiaries assigned to plans that used no authorization restrictions will not have differences in spending across the two simulations.

[18] This incorporates both the days supply consumed and the transacted price. One wrinkle is that the true net price of procuring a drug includes rebates offered by manufacturers; as Kakani et al. (2020) show, these rebates are economically meaningful. However, we do not have access to the data required to do such an adjustment. Our results therefore overstate the total spending reductions.

V. Administrative Cost Burdens

As we discuss in Section I.B, when introducing prior authorization we need to trade off the spending reductions we estimate above against the paperwork costs of bureaucracy. In this section, we attempt to estimate those costs.

Unfortunately, we have no data on the steps in the bureaucratic process of prior authorization. Therefore, unlike Dunn et al. (2021), we cannot compute rejections rates, nor can we estimate the costs of compliance. We therefore take an alternative approach, wherein we calibrate relevant parameters (per-application costs, and rejection rates) and combine them with our demand system estimates to estimate the total paperwork burden.

Specifically, we assume that, for any beneficiary-drug pair, authorization must be received once in a given year if it is required and if the patient wishes to obtain the drug.[19] We assume that making a request incurs some constant joint cost to both the requesting physician as well as the insurer, which we call a. The number of requests is also unobserved. We assume that any patient we observe taking a restricted drug must have made an authorization request; we also assume that there are some who made a request but were rejected at a constant rate r, and thus we do not observe them taking the drug. If we observe N patients taking the drug, with a rejection rate r, there will have been $\frac{N}{1-r}$ requests.[20] Therefore, the total administrative costs are $\frac{aN}{1-r}$.

We calibrate a and r from prior studies in the health policy literature. For N, we simply use our demand system to estimate the number of beneficiaries choosing restricted drugs in the status quo simulation, summed across classes. The average beneficiary chooses 0.316 restricted drugs per year.

Calibrating Application Costs and Rejection Rates

When considering costs of prior authorization, there are two parties who incur costs for each authorization request: Medical providers, who need to submit requests, and insurers, who need to process and respond to them. We draw from case studies and industry reports to calibrate a measure of both costs. There are three sources that we are aware of that have estimated provider-side costs: Raper et al. (2010), who do so for a single HIV clinic in Alabama between March 2006 and February 2008, Bukstein et al. (2006), who do the same in an allergist clinic in Madison, WI, in August to October 2003, and Council for Affordable Quality Healthcare (2014), who field an annual survey of providers. Raper et al. compute both direct and opportunity costs (i.e., the revenue that could have been

[19] In general, authorization is required once per treatment course, but this is heterogeneous across drugs and unobserved. In practice, this means authorization may be required less or more than once a year.

[20] We abstract from repeat interactions between the requesting physician and the insurer. Additionally, we are abstracting from real heterogeneity in both the costs and rejection rates associated with making requests for different kinds of drugs.

earned by the nurse practitioner who filed the request if he or she instead saw patients for standard office visits), whereas Bukstein et al. and CAQH only measure direct costs.[21] Raper et al.'s direct estimate (which includes nurse practitioner and administrative staff time at their wage levels, as well as materials cost) is \$14.24, whereas their opportunity cost estimate comes out to \$27.35. Bukstein et al. (2006) estimate direct costs only, estimating them at \$17.77 per request. The CAQH estimate ranges from year to year, but their 2013 report (the earliest we were able to find, using a survey fielded in 2012, report completed in 2014) estimates direct costs at \$18.53.[22] To compute insurer costs, we use a similar insurer-facing survey from CAQH, which estimated manual processing costs of \$3.95 for insurers in 2012.[23] Added together, that gives us total cost per application estimates of \$18.19, \$21.72, \$22.48, and \$31.30. We also experiment with a handful of more extreme values: \$50, \$100, and \$200.

The literature provides many more estimates of prior authorization request rejection rates, although not all of them are directly comparable, and none precisely get at the exact quantity of interest–the number of (unobserved) requests per (observed) successful fill. Nonetheless, we take a handful of measures from this literature. We report the rates from the universe of studies we found in Appendix Table A2. Unfortunately, none of them are directly comparable to our setting. The majority cover either a single, potentially unrepresentative clinic, or have extensive coverage that includes unrelated products (e.g. hospital services). We use five values: 1.5%, 4%, 7.5%, and 15%, which cover the range of estimates found in the literature, as well as 0%.

Computing Net Fiscal Savings from Authorization Restrictions

With these calibrated values in hand, we can compute the total administrative burden generated by prior authorization. As in Section IV, we consider the burden generated by moving between the historical status quo and a counterfactual world in which prior authorization was removed but exclusion left intact. For every pair of calibrated values of a and r, we report in Table 6 the amount of administrative costs, in dollars, induced for each dollar saved in drug spending. If this ratio is below 1, then the total administrative cost burden is below the drug spending savings generated by prior authorization restriction policies, and such policies will have generated net fiscal savings. If this ratio is above 1, then prior authorization will have generated net fiscal costs.

[21] In their paper, Raper et al. incorrectly add these two cost measures together, which double-counts the value of the nurse practitioner's time.

[22] We use their estimate for manually-submitted requests. Costs for doing so through an IT system were estimated at only \$5.20, but the majority of requests (110 million out of 130 million) were filed manually. Their cost estimates for manual filing decreased in later reports, with \$14.07 for calendar year 2013, \$7.17 for 2014, and \$7.50 for 2015.

[23] Manual insurer-facing costs are stable across time in the CAQH survey and never exceed \$3.95 per request. Bukstein et al. (2006) claim that insurer costs for processing nonformulary drug requests are \$20-\$25 based on personal communication with a pharmacy manager at a major PBM, but since the methodology of estimating these costs is unclear we discard this estimate.

Unsurprisingly, at higher calibrated values of a and r, administrative costs per dollar saved are higher. However, prior authorization policies are still estimated to generate positive net fiscal savings unless we assume that unit per-application costs are roughly \$200 or more, well above those previously documented. Our preferred calibrated measure of a is \$22.48, which adds both CAQH-measured sources of costs. For this measure, across our range of rejection rate calibrations, 12 to 14 cents of administrative costs are induced for each dollar saved; in total, this is \$7.3 to \$8.5 per beneficiary-year, an order of magnitude below the savings generated; the net fiscal savings

Table 6—: Net Fiscal Savings From Authorization Restrictions

		Request Rejection Rate				
		0%	1.5%	4%	7.5%	15%
Paperwork Cost	\$18.19	0.09	0.10	0.10	0.10	0.11
	\$21.72	0.11	0.11	0.12	0.12	0.13
	\$22.48	0.12	0.12	**0.12**	0.13	0.14
	\$31.30	0.16	0.16	0.17	0.18	0.19
	\$50	0.26	0.26	0.27	0.28	0.31
	\$100	0.52	0.53	0.54	0.56	0.61
	\$200	1.04	1.05	1.08	1.12	1.22

Note: This table reports estimates of the ratio of administrative costs of prior authorization to the reductions in drug spending induced by it, from the historical authorization restriction regimes implemented in Medicare Part D. Each cell represents the estimate under a calibrated set of values for the application cost a and rejection rate r. Ratios below 1 indicate net fiscal savings, while ratios above 1 indicate net costs.

While these results are applied to the sum across all therapeutic classes, there are good reasons to think they should differ across classes. As we discuss in Section I.B, prior authorization is likely to be most beneficial when the savings from deterring marginal users are relatively large, and when the number of inframarginal users is relatively small. The extent to which this is true varies across classes. In Figure 3, we plot, for each therapeutic class, the total dollars saved per inframarginal user, i.e., in our above terminology, $\frac{\Delta \text{Spend}}{N}$. We can then compare this value to various calibrations of $\frac{a}{1-r}$ to determine net fiscal savings for a specific class. The range of calibrated values is given by the two red lines in the figure (the left representing $a = 18.19$ and $r = 0$, the right representing $a = 31.30$ and $r = .15$), and the confidence intervals are given by the black brackets.

We can see that, for many classes, we can reject that net fiscal savings are zero. The class with the largest savings per inframarginal user is the class of biologic response modifiers, a class where very few beneficiaries receive any drug at all, and where each individual drug is quite expensive,

We can now use this to bound how large patient valuations must be for authorization restrictions to be beneficial for welfare. Substituting calibrated values

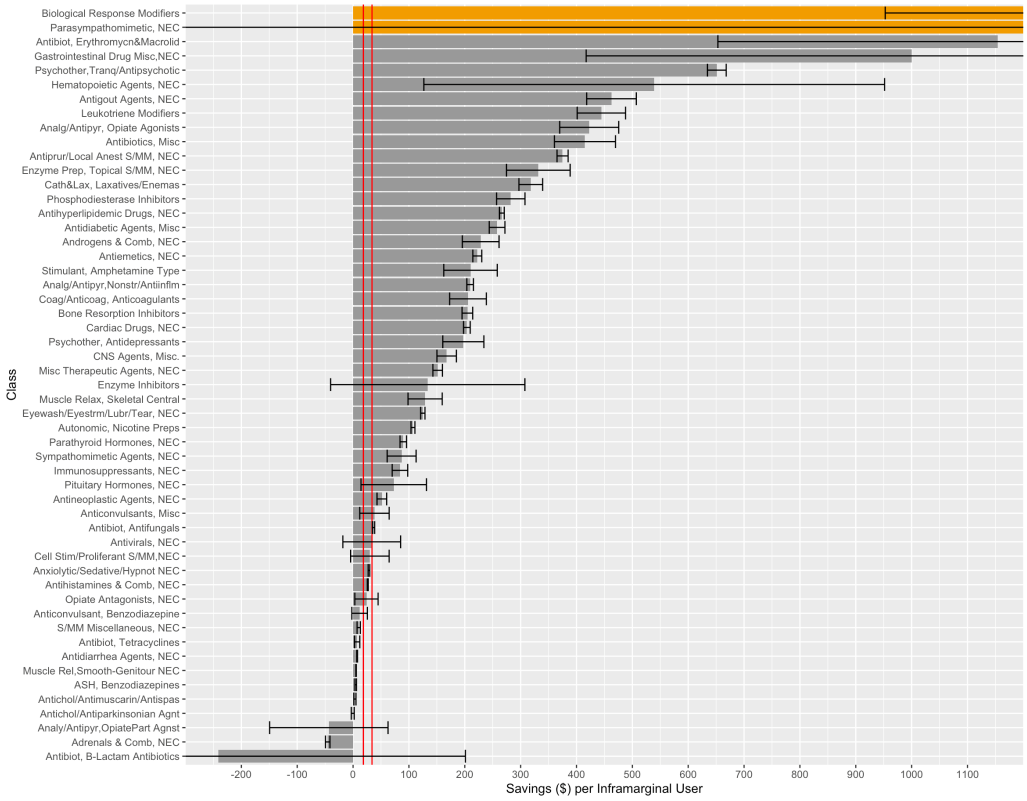


Figure 3. : Ratio of Drug Cost Reduction to Administrative Cost Burden by Class

Note: This table reports estimates of the ratio of dollars saved from prior authorization to the number of inframarginal drug users, for each therapeutic class. Black brackets represent 95% confidence intervals, and the two red lines represent bounds of our calibrated value for $\frac{a}{1-r}$; the left line representing $a = 18.19$ and $r = 0$, the right representing $a = 31.30$ and $r = .15$.

into Equation **??**, we have

$$\sum_i \Delta v_i D_i^H(0) D_i^L(0) + \sum_i v_i^H D_i^H(0) D_i^0(1) < \$53.7$$

as our condition. Our prior estimates showed that prior authorization deterred 0.08 restricted drug uses, with 44.5% going to another drug and 55.5% going to no drug. This requirement therefore becomes:

$$0.08 \times (\Delta v_i \times 0.445 + v_i \times 0.555) < \$53.7$$

or $\Delta v_i \times 0.445 + v_i \times 0.555 < \$671.25$. That is, the average beneficiary who marginally substitutes must value the average deterred drug by \$671.25 per year relative to their alternative. Given that paperwork costs are roughly 12% savings, this means that marginal beneficiaries must value this consumption at 88% of the

procurement cost, a substantial bar for top-up consumption. In our final section below, we try to assess the value of this deterred consumption.

VI. Are Prior Authorization Restrictions Efficient?

Our results in Section V suggest that, as long as equilibrium application-per-approved drug rates/paperwork costs are low enough, authorization restrictions need to destroy a significant amount of drug consumption value for their use to be harmful. While we cannot test this directly, we take a number of approaches that test for whether the effects of authorization restrictions are relatively inefficient. We explore this on two dimensions: First, we explore heterogeneity in the effects of prior authorization policies based on characteristics of patients that should not be expected to have differential effects. Second, we explore health effects in a specific case, anticoagulation, where prior authorization is common and health effects are easily measured.[24]

A. Heterogeneous Effects by Patient Characteristics

An optimal implementation of a screening mechanism should, ideally, screen on relevant characteristics and *not* screen on irrelevant characteristics. We test for the latter by asking whether the effects of authorization restrictions are stronger by gender and race. In this context, this is a particular worry. Authorization is intermediated by medical providers and their willingness to exert effort on behalf of their patients. Prior research has shown that providers exert less effort for women and non-white patients, meaning that such patients may be more affected by provider-facing bureaucracy.

We test this by re-estimating our model from Equation 1 separately for white vs. non-white patients, and male vs. female patients. This is effectively the same as allowing all of our covariates to interact with race or gender (not just the treatment, but also all drug-region-year and plan-region-year fixed effects). Life expectancy is different across race and gender, which may affect the composition of the two samples. We therefore weight observations in our regressions for male and non-white patients to match female and white patients, i.e., proportional to $\frac{P[X_i|\text{Female}]}{P[X_i|\text{Male}]}$ for our male regression and proportional to $\frac{P[X_i|\text{White}]}{P[X_i|\text{Non-White}]}$ for our non-white regression, where X_i are cells of age (in 5-year buckets) and health status (in Elixhauser quintiles).

We report the results in Table 7. Table 7a displays our results segmented by gender and Table 7b displays our results segmented by race. We find level differences in the effects of restrictions by gender, with women facing a (statistically significantly) larger burden than men both in terms of spending and in terms of

[24]This draft is incomplete. We are currently exploring four more angles: 1) Heterogeneous effects by value of drug; 2) Heterogeneous effects by provider characteristics; 3) The extent to which deterred consumption is by patients who are less likely to stay on the drug, as an indicator of their match value; 4) Effects on mortality.

extensive margin drug take-up. While there are some level differences in the extensive margin of use for nonwhite patients, these are not statistically significant.

However, it is important to note that baseline utilization in the absence of restrictions also differs across these groups. If we rescale the treatment effects by baseline utilization (i.e., compute the percent change induced by authorization restrictions rather than the levels), we see a surprising reversal. A substantial racial gap emerges, with extensive margin effects for white patients being a reduction of 20%, whereas nonwhite patients see reductions of 25%. Conversely, the gender gap shrinks, considerably such that we can no longer reject differences by gender, with both genders seeing roughly 22% reductions in extensive margin use. This is due to the fact that non-white patients have substantially lower baseline drug use than white patients, whereas women have substantially *higher* baseline use than men. There is no clear guideline as to whether absolute or relative effects should be given more credence when considering disparate effects (especially if the baseline already suffers from disparity), but it is clear that authorization restrictions have some undesirable disparate impacts.

B. Health Effects

One of the clearest ways that authorization restrictions could harm patients is if their reduction in drug utilization harms patient health. The American Medical Association's public documents on prior authorization make strong claims about potential health harms, claiming restrictions might lead to "hospitalization, disability, or even death," although these claims are based on a survey of physicians' opinions rather than measured outcomes.[25] Our primary research design does not permit measurements of health effects, since health is defined at the patient level, whereas our research design is identified at the patient-drug level. To estimate health effects, we either need drug-specific measures of health, or an alternative design that induces variation at the patient level. We start by using the former.

As a measure of drug-specific health, we measure the effects of prior authorization on a specific class of drug: Oral anticoagulants, often referred to as blood thinners. Anticoagulants reduce the extent of blood clotting, therefore reducing the risk of strokes (blood clots that occur in the brain), as well as other clot-driven health issues such as heart attacks and pulmonary embolisms, and are used regularly over a long period of time. The standard anticoagulant from 1954 until the 2010s was warfarin, which, by the time our sample period had begun, existed primarily as a cheap generic, costing approximately $0.30 per pill. In the 2010s, however, a series of drugs called non-Vitamin K oral anticoagulants (NOACs) were approved by the FDA and introduced into usage. There are two main advantages of these new drugs over warfarin: required dose varies less, so there is less need for frequent monitoring of blood clotting, and there are fewer food and drug-interactions. By 2015, NOACs represented around one-eighth of

[25] c.f. `https://www.fixpriorauth.org/patients`.

Table 7—: Regressions with Patient Heterogeneity

(a) Gender

	Spending		% Ever filled	
	Female	Male	Female	Male
$\text{Auth}^{\text{Assigned}}$	-0.844	-0.659	-0.089	-0.066
	(0.042)	(0.044)	(0.003)	(0.003)
$\text{Excluded}^{\text{Assigned}}$	-1.068	-0.827	-0.115	-0.082
	(0.034)	(0.034)	(0.003)	(0.003)
Auth^{sub}	0.186	0.111	0.042	0.031
	(0.037)	(0.040)	(0.003)	(0.003)
$\text{Excluded}^{\text{sub}}$	0.560	0.463	0.122	0.082
	(0.049)	(0.051)	(0.006)	(0.005)
% Effect of $\text{Auth}^{\text{Assigned}}$	-28.4	-23.4	-12.5	-12.2
	(1.4)	(1.5)	(0.5)	(0.5)
Control Mean	3.553	3.557	1.087	0.827
Reweighted Control Mean	3.744	3.429	0.363	0.264
N (plan-drug-market-years)	2,141,993	2,141,993	2,141,993	2,141,993

(b) Race

	Spending		% Ever filled	
	White	Non-White	White	Non-White
$\text{Auth}^{\text{Assigned}}$	-0.764	-0.774	-0.076	-0.084
	(0.041)	(0.047)	(0.003)	(0.004)
$\text{Excluded}^{\text{Assigned}}$	-0.964	-0.973	-0.095	-0.109
	(0.030)	(0.041)	(0.002)	(0.004)
Auth^{sub}	0.194	0.094	0.039	0.035
	(0.032)	(0.044)	(0.002)	(0.004)
$\text{Excluded}^{\text{sub}}$	0.530	0.504	0.100	0.112
	(0.046)	(0.058)	(0.005)	(0.008)
% Effect of $\text{Auth}^{\text{Assigned}}$	-24.9	-29.0	-11.6	-13.6
	(1.3)	(1.7)	(0.4)	(0.6)
Control Mean	3.744	3.260	0.999	0.947
Reweighted Control Mean	3.887	3.178	0.335	0.301
N (plan-drug-market-years)	2,141,993	2,141,993	2,141,993	2,141,993

Note: This table presents estimates from regressions of utilization measures for each demographic group on prior authorization and exclusion of a given drug and other drugs within the class.

anticoagulant prescriptions, but two-thirds of spending (see Figures A2 and A3). Total anticoagulation spending rose substantially over this period, exemplary of

the hypothesis by Chandra and Skinner (2012) that technological change drives spending increases.

We restrict to a subsample of individuals with a medical history of atrial fibrillation, deep vein thrombosis, or pulmonary embolism, the typical conditions treated by anticoagulants. Most plans put authorization restrictions on all NOACs when they are available, or cover them without restriction. Therefore we run the regression

$$Y_i = \beta \text{AuthAllNOACs}_i + \gamma \text{OtherFormulary}_i + \delta_{m(i)} + \epsilon_i$$

where OtherFormulary_i includes any formulary other than restricting all NOACs or unrestrictedly covering all NOACs (all plans fully cover warfarin), so β measures the difference between restriction and unrestriction. We omit plan fixed effects since we estimate this regression for a single set of drugs at a time, and thus plan identifiers are colinear with the treatment.

We estimate β values for utilization of any anticoagulant, utilization of warfarin, and utilization of any NOAC, and report them in Table 8. As in our prior analysis, authorization restrictions reduce overall spending on anticoagulants, but do not significantly imapct the total use of anticoagulants, instead reallocating around 25% of patients from NOACs to warfarin. We then estimate the effect on health outcomes: The probability of death during the year, the probability of a stroke, and the probability of a bleeding event. We report the results from these regressions in Table 9. We find small, insignificant effects for all three, and can reject large effect sizes in either direction. This is consistent with meta-analyses in the medical literature finding limited differences in clinical outcomes between warfarin and NOACs, with most of the differences coming through side effects and ease of use.

Table 8—: Effects of anticoagulant prior authorization restrictions on outcomes

	Spending			Any prescription		
	All	NOACs	Warfarin	All	NOACs	Warfarin
All NOACs PA	-16.6	-18.3	1.7	-0.0003	-0.0097	0.0069
	(6.41)	(6.61)	(0.70)	(0.0033)	(0.0028)	(0.0035)
Other restrictions	-16.1	-12.4	-3.7	-0.0011	-0.0100	0.0058
	(5.51)	(5.14)	(2.51)	(0.0086)	(0.0035)	(0.0085)
R^2	0.026	0.032	0.019	0.014	0.030	0.024
Mean	111.585	77.433	34.152	0.291	0.043	0.260
N bene years	134,182					
N market years	160					

Note: This table presents estimates from a set of regressions of health outcome of individual i on dummies for whether their plan has prior authorization restrictions on all NOACs. Regressions include market-year fixed effects.

Table 9—: Effects of anticoagulant prior authorization restrictions on outcomes

	Died	Stroke	Bleed
All NOACs under prior auth	-0.00032	0.00067	-0.00124
	(0.00090)	(0.00143)	(0.00172)
Other restrictions on NOACs	-0.00152	0.00569	0.00480
	(0.00189)	(0.00275)	(0.00497)
R^2	0.002	0.002	0.002
Mean of health variable	0.014	0.032	0.054
Number of bene years		134,182	
Number of market years		160	

Note: This table presents estimates from a set of regressions of health outcome of individual i on dummies for whether their plan has prior authorization restrictions on all NOACs. Regressions include market-year fixed effects.

While this is one example, we think it is generally representative: In oral anticoagulants, as in many other classes, substitution is largely to another drug, rather than to no drugs. While, again, this is not fully representative of patient value for the change in their drug regimen, these results suggest that, in contrast to recent work finding that cost-sharing for drugs in Medicare Part D increases patient mortality (Chandra et al., 2021), prior authorization restrictions are a managed care tool that does not have such an effect.

C. Revealed Preference Approaches

In this section we use two complementary revealed preferences approaches to estimate the consumption value of drug utilization deterred by prior authorization restrictions. First, we estimate the copayment reduction that would be required to deter the same quantity of drug consumption as prior authorization. Second, we use our dollar estimates of physician administrative costs to infer the value of consumption as perceived by the physician.

Since our main sample includes only beneficiaries in the LIS program who do not face drug copayments, we cannot use this population to estimate elasticity of demand with respect to copayments. Instead we use the set of beneficiaries who go from facing drug copayments to no copayments when they transition into the LIS program, of whom we see 62,785. We observe each of these beneficiaries one year prior to their transition, the year of transition and one year post-transition. We estimate the following regression:

$$\log\left(\mathbb{E}[Y_{idt}]\right) = \beta P_{dt} \times NotLIS_{it} + \alpha_i + \gamma_{dmt} + \epsilon_{it}$$

Drug price P_{dt} is the average copayment per script across all users of the drug in year t. This means we are using only the variation that comes from transitioning into LIS, and not between-plan differences in copayments for a given drug.

$NotLIS_{it}$ is a dummy equal to 1 if the individual has not yet transitioned into LIS. To make this estimate comparable to our estimates of the effect of prior authorization, we weight observations by $(\text{Var}[\text{Auth}_{id}|d,m])/P_{dt}$, the variance of authorization restriction status for a given drug in a given market across beneficiaries. β estimates the price semi-elasticity of demand.

Table 10 shows the results from this regression. The semi- elasticity of demand with respect to copayment is approximately 0.4. Prior authorization on a focal drug deters 24.8% of consumption. Assuming a demand curve with constant price semi-elasticity of demand, this implies that imposing prior authorization on a drug is equivalent to raising its price by $95.

We impose this constant semi-elasticity onto our demand estimation sample. In this sample, we estimate that there are roughly 0.32 drug users per capita under no prior authorization. Therefore, we calibrate the relevant demand curve as $D(P) = 0.32e^{-0.38p}$.

We then construct consumer surplus loss in two ways. First, we assume that the 24.8% of marginal drug users are drawn from those with the lowest value for the drug, as we would expect if prior authorization restrictions perfectly screen on value. To compute this, we have

$$\Delta CS^{\text{Perfect}} = -\int_{75.2\%}^{100\%} D^{-1}(Q)dQ$$

where Q represents the share of patients taking the drug. This only applies if only the lowest-value patients are screened out. However, from earlier subsections, we know that screening is not especially well-targeted. Therefore, we conduct an additional measure where screening is random. In this case, a random 24.8% chunk is taken out of the full consumer surplus in expectation. To compute this, we have

$$\Delta CS^{\text{Random}} = -0.248\int_{0\%}^{100\%} D^{-1}(Q)dQ$$

Unsurprisingly, these two measures give us radically different results, as shown in Table 11. Under perfect screening, consumer surplus loss is small; only $3.65 per beneficiary-year. Under random screening, the loss is substantial, nearly $26.67. Worse-than-random screening (as found by e.g. Deshpande and Li (2019)) would result in even greater consumer surplus loss. This is of the same order of magnitude as the net fiscal savings of applying prior authorization, albeit smaller.

To estimate value of consumption loss using physician revealed preference, consider a physician deciding whether to prescribe restricted drug d to physician i. Physician utility is as follows

$$u_{id} = \theta(v_{id} - v_{i,-d}) - a$$

where v_{id} is patient valuation of the focal drug is v_{id}, and $v_{i,-d} = \max_{k\neq d} v_{ik}$ is

Table 10—: Effect of drug copayments on utilization for beneficiaries transitioning into the LIS program

	Prescription Count	Ever Filled
β (Price semi-elasticity)	-0.0038	-0.0038
	(0.0001)	(0.0001)
N	229.1m	229.1m
Number of benes	62,785	62,785

patient valuation of the next best option, θ is the weight the physician places on patient preferences and a is administrative cost. We use the highest physician-facing cost that we calibrate, which is \$27.35. With such costs, deterring 24.8% of prescriptions implies that the administrative cost semi-elasticity of prescription is roughly 0.91. We focus on the case where physicians are perfectly altruistic, i.e., $\theta = 1$, and again consider both perfect screening and random screening. If screening is perfect, we can compute a nonparametric upper bound, i.e., the physician would only fail to prescribe if $v_{id} - v_{i,-d} < a$, and so $\Delta Q \times a$ is a bound on the total consumer surplus loss. This is equal to \$2.19 per beneficiary-year.

Under random screening, we cannot compute a nonparametric bound, since we cannot make guarantees about the demand curve beyond those who are marginal. We assume two demand projections: Linear, and constant cost semi-elasticity. As with our patient approach, to compute surplus changes under perfect screening, we integrate under the projected demand curve for the 24.8% of patients with the lowest value for the drug. Under random screening, we integrate under the full demand curve, and multiply by 0.248.

For perfect screening, our parametric approaches unsurprisingly compute losses that are below the nonparametric upper bound. As we allow for random screening, losses become larger, both hovering around \$8.75 per beneficiary-year. This is substantially smaller than both net fiscal savings and our estimates from the patient approach; the latter may reflect our likely-incorrect assumption that physicians are perfect agents. In practice, if we believe all parts of the physician approach other than the assumptions on θ, and we believed prior authorization was welfare-reducing, we would need to believe that θ is at most 0.02 (under perfect screening) or at most 0.16 (under random screening).

VII. Conclusion

We have found that prior authorization policies, applied to drugs in Medicare Part D, lower program spending by an amount that exceeds the administrative costs of prior auth to physicians. While the effect of prior authorization is not necessarily efficiently allocated, with potentially larger deterrence of use for non-white and female patients, the reduction in consumption seems to have a minimal

Table 11—: Estimates of consumer surplus loss via revealed preference

	Physician			Patient
	Nonparam. Up. Bound	Linear	Cons. Semi-Elasticity	Cons. Semi-Elasticity
Perfect screening	2.19	1.09	0.93	3.65
Random screening	-	8.75	8.75	26.67
Semi-elasticity		0.91		0.04

impact on patient health.

To conclude, we would like to emphasize two broader points that arise from our work. First, although these policies reduce net social costs, they *raise* costs for physician and other health care providers, by increasing their paperwork burdens. These policies are Kaldor-Hicks efficient in the sense that providers could be transferred a portion of the savings to be made at least indifferent between being the stewards of these policies and not. However, much of the gains are realized by insurers, and, in particular in our setting, drug insurers who have no direct interaction with providers. Finding a way to efficiently share the gains with providers is a serious political economy issue, particularly in the light of the American Medical Association's strong opposition to paperwork burdens.

Second, our results have important implications for the broader discourse around international health care spending comparisons and U.S. health care reform. This paper shows that, although managed care rationing mechanisms introduce seemingly-wasteful administrative costs, they carry the net benefit of reducing overall costs. However, because the costs are largely incurred by providers and insurers, they show up in accounting measures of costs. In contrast, using queueing more aggressively, as is the case in other OECD health systems, imposes costs on patients that are not captured in accounting measures. More research is needed to both characterize the costs and benefits of other sources of administrative cost burden, as well as to compare how other rationing mechanisms induce administrative costs, both those that show up in accounting data and those that do not.

REFERENCES

Akerlof, George A., "The Economics of "Tagging" as Applied to the Optimal Income Tax, Welfare Programs, and Manpower Planning," *American Economic Review*, 1978, *68* (1), 8–19.

Alpert, Abby E., Sarah E. Dykstra, and Mireille Jacobson, "How Do Prescription Drug Monitoring Programs Reduce Opioid Prescribing? The Role of Hassle Costs versus Information," 2020. NBER Working Paper No. 27584.

American Medical Association, "2017 AMA Prior Authorization Physician Survey," 2017.

America's Health Insurance Providers, "Key Results of Industry Survey on Prior Authorization," 2020. https://www.ahip.org/wp-content/uploads/Prior-Authorization-Survey-Results.pdf.

Archer, Diane, "Medicare Is More Efficient Than Private Insurance," *Health Affairs Blog*, 2011. https://www.healthaffairs.org/do/10.1377/hblog20110920.013390/full/.

Baicker, Katherine, Sendhil Mullainathan, and Joshua Schwartzstein, "Behavioral Hazard in Health Insurance," *Quarterly Journal of Economics*, 2015, *130*, 1623–1667.

Bajari, Patrick, Stephanie Houghton, and Steven Tadelis, "Bidding for Incomplete Contracts: An Empirical Analysis of Adaptation Costs," *American Economic Review*, 2014, *104* (4), 1288–1319.

Berry, Steven and Philip Haile, "Identification in Differentiated Products Markets," *Annual Review of Economics*, 2016, *8*, 27–52.

Berry, Steven T., "Estimating Discrete-Choice Models of Product Differentiation," *RAND Journal of Economics*, 1994, pp. 242–262.

Birdsall, Ashley D., Ashley M. Kappenman, Bryce T. Covey, and Matthew H. Rim, "Implementation and impact assessment of integrated electronic prior authorization in an academic health system," *Journal of the American Pharmacists Association*, 2020, *60* (4), e93–e99.

Brot-Goldberg, Zarek C., Amitabh Chandra, Benjamin R. Handel, and Jonathan T. Kolstad, "What Does a Deductible Do? The Impact of Cost-Sharing on Health Care Prices, Quantities, and Spending Dynamics," *Quarterly Journal of Economics*, 2017, *132* (3), 1261–1318.

_ **and Mathijs de Vaan**, "Intermediation and Vertical Integration in the Market for Surgeons," 2018.

_ **, Timothy Layton, Boris Vabson, and Adelina Yanyue Wang**, "The Behavioral Foundations of Default Effects: Theory and Evidence from Medicare Part D," 2021. NBER Working Paper No. 28331.

Bukstein, Don A., Greeta A. Cherayil, Adam D. Gepner, Allan T. Luskin, J. Brent Kooistra, and Reid M. Olson, "The economic burden associated with prior authorizations in an allergist office," *Allergy & Asthma Proceedings*, 2006, *27* (2), 119–122.

Cabral, Marika and Marcus Dillender, "Gender Differences in Medical Evaluations: Evidence from Randomly Assigned Doctors," 2022. NBER Working Paper No. 29541.

Casalino, Lawrence P., Sean Nicholson, David N. Gans, Terry Hammons, Dante Morra, Theodore Karrison, and Wendy Levinson, "What Does It Cost Physician Practices To Interact With Health Insurance Plans?," *Health Affairs*, 2009, *28* (4), w533–w543.

Chandra, Amitabh and Jonathan Skinner, "Technology Growth and Expenditure Growth in Health Care," *Journal of Economic Literature*, 2012, *50* (3), 645–680.

_ , Evan Flack, and Ziad Obermeyer, "The Health Costs of Cost-Sharing," 2021. NBER Working Paper No. 28439.

Correia, Sergio, Paulo Guimarães, and Thomas Zylkin, "Fast Poisson Estimation with High-Dimensional Fixed Effects," *Review of Economics and Statistics*, 2020, *20* (1), 95–115.

Council for Affordable Quality Healthcare, "2013 U.S. Healthcare Efficiency Index: Electronic Administrative Transaction Adoption and Savings," 2014.

Cutler, David M., "Reducing Administrative Costs in U.S. Health Care," 2020. The Hamilton Project.

_ and Dan P. Ly, "The (Paper)Work of Medicine: Understanding International Medical Costs," *Journal of Economic Perspectives*, 2011, *25* (2), 3–25.

_ , Elizabeth Wikler, and Peter Basch, "Reducing Administrative Costs and Improving the Health Care System," *New England Journal of Medicine*, 2012, pp. 1875–1878.

de Chaisemartin, Clement and Xavier D'Haultfœuille, "Two-Way Fixed Effects Estimators with Heterogeneous Treatment Effects," *American Economic Review*, 2020, *110* (9), 2964–96.

Deshpande, Manasi and Yue Li, "Who Is Screened Out? Application Costs and the Targeting of Disability Programs," *American Economic Journal: Economic Policy*, 2019, *11* (4), 213–248.

Dillender, Marcus, "What Happens When the Insurer Can Say No? Assessing Prior Authorization as a Tool to Prevent High-Risk Prescriptions and to Lower Costs," *Journal of Public Economics*, 2018, *165* (1), 170–200.

Dranove, David and Kathryn E. Spier, "A Theory of Utilization Review," *Contributions in Economic Analysis & Policy*, 2003, *2* (1), 1–19.

Dubois, Pierre, Ashvin Gandhi, and Shoshana Vasserman, "Bargaining and international reference pricing in the pharmaceutical industry," 2019.

Dunn, Abe, Joshua D. Gottlieb, Adam Shapiro, Daniel J. Sonnenstuhl, and Pietro Tebaldi, "A Denial a Day Keeps the Doctor Away," 2021.

_ , _ , and Adam Hale Shapiro, "Administration Above Administrators: The Changing Technology of Health Care Management," *AEA Papers and Proceedings*, 2020, *110*, 274–278.

Eliason, Paul, Riley League, Jetson Leder-Luis, Ryan C. McDevitt, and James W. Roberts, "Ambulance Taxis: The Impact of Regulation and Litigation on Health Care Fraud," 2021. NBER Working Paper No. 29491.

Finkelstein, Amy and Matthew J. Notowidigdo, "Take-up and Targeting: Experimental Evidence from SNAP," *The Quarterly Journal of Economics*, 2019, *134* (3), 1505–1556.

Frakt, Austin, "Is Medicare for All the Answer to Sky-High Administrative Costs?," *The New York Times*, 2018.

Ganapati, Sharat and Rebecca McKibbin, "Non-tariff barriers and bargaining in generic and off-patent pharmaceuticals," 2019.

Gennaioli, Nicola, Rafael La Porta, Florencio Lopez de Silanes, and Andrei Shleifer, "Trust and Insurance Contracts," 2020. NBER Working Paper No. 27189.

Glied, Sherry, "Managed Care," in "Handbook of Health Economics," Vol. 1, Elsevier, 2000, pp. 707–753.

Gottlieb, Joshua D., Adam Hale Shapiro, and Abe Dunn, "The Complexity Of Billing And Paying For Physician Care," *Health Affairs*, 2018, *37* (4), 619–626.

Guimarães, Paulo, Octávio Figueiredo, and Douglas Woodward, "A Tractable Approach to the Firm Location Decision Problem," *Review of Economics and Statistics*, 2003, *85* (1), 201–204.

Heinrich, Carolyn J., "The Bite of Administrative Burden: A Theoretical and Empirical Investigation," *Journal of Public Administration Research and Theory*, 2016, *26* (3), 403–420.

Herd, Pamela and Donald P. Moynihan, *Administrative Burden: Policymaking By Other Means*, Russell Sage Foundation, 2018.

Himmelstein, David U., Terry Campbell, and Steffie Woolhandler, "Health Care Administrative Costs in the United States and Canada, 2017," *Annals of Internal Medicine*, 2020, *172* (2), 134–142.

Ho, Kate and Ariel Pakes, "Hospital Choices, Hospital Prices and Financial Incentives to Physicians," *American Economic Review*, 2014, *104* (12), 3841–3884.

Homonoff, Tatiana and Jason Somerville, "Program Recertification Costs: Evidence from SNAP," *American Economic Journal: Economic Policy*, forthcoming.

Kahan, Natan R., David P. Chinitz, Dan-Andrei Waitman, and Ernesto Kahan, "When gatekeepers meet the sentinel: the impact of a prior authorization requirement for cefuroxime on the prescribing behaviour of community-based physicians," *British Journal of Clinical Pharmacology*, 2006, *61* (3), 341–344.

Kakani, Pragya, Michael Chernew, and Amitabh Chandra, "Rebates in the Pharmaceutical Industry: Evidence from Medicines Sold in Retail Pharmacies in the U.S.," 2020. NBER Working Paper No. 26846.

LaPensee, Kenneth, "Analysis of a Prescription Drug Prior Authorization Program in a Medicaid Health Maintenance Organization," *Journal of Managed Care Pharmacy*, 2003, *9* (1), 36–44.

Lieber, Ethan M.J. and Lee M. Lockwood, "Targeting with In-Kind Transfers: Evidence from Medicaid Home Care," *American Economic Review*, 2019, *109* (4), 1461–1485.

Nichols, Albert L. and Richard J. Zeckhauser, "Targeting Transfers through Restrictions on Recipients," *American Economic Review*, 1982, *72* (2), 372–377.

Pauly, Mark V., "The Economics of Moral Hazard: Comment," *American Economic Review*, 1968, *58* (3), 531–537.

Pozen, Alexis and David M. Cutler, "Medical Spending Differences in the United States and Canada: The Role of Prices, Procedures, and Administrative Expenses," *Inquiry*, 2010, *47* (2), 124–134.

Raper, James L., James H. Willig, Hui-Yi Lin, Jeroan J. Allison, M. Bennet Broner, Michael J. Mugavero, and Michael S. Saag, "Uncompensated Medical Provider Costs Associated with Prior Authorization for Prescription Medications in an HIV Clinic," *Clinical Infectious Diseases*, 2010, *51* (6), 718–724.

Sarig, Oren, "Pharmaceutical demand response to utilization management," 2020.

Schwartz, Aaron L., Troyen A. Brennan, Dorothea J. Verbrugge, and Joseph P. Newhouse, "Measuring the Scope of Prior Authorization Policies:

Applying Private Insurer Rules to Medicare Part B," *JAMA Health Forum*, 2021, *2* (5), e210859–e210859.

Seabury, Seth A., Darius N. Lakdawalla, Deborah Walter, John Hayes, Thomas Gustafson, Anshu Shrestha, and Dana P. Goldman, "Patient Outcomes and Cost Effects of Medicaid Formulary Restrictions on Antidepressants," 2014, *17* (2), 153–168.

Shepard, Mark and Myles Wagner, "Reducing Ordeals through Automatic Enrollment: Evidence from a Subsidized Health Insurance Exchange," 2021.

Shi, Maggie, "The Costs and Benefits of Monitoring Providers: Evidence from Medicare Audits," 2021.

Song, Zirui, Yunan Ji, Dana G. Safran, and Michael E. Chernew, "Health Care Spending, Utilization, and Quality 8 Years into Global Payment," *New England Journal of Medicine*, 2019, *381* (3), 252–263.

Sood, Neeraj, Zachary Wagner, Peter J. Huckfeldt, and Amelia M. Haviland, "Price Shopping in Consumer Directed Health Plans," *Forum for Health Economics & Policy*, 2013, *16* (1), 35–53.

U.S. Department of Health and Human Services Office of Inspector General, "Medicare Advantage Appeal Outcomes and Audit Findings Raise Concerns About Service and Payment Denials," 2018. Report OEI-09-16-00410.

Waldinger, Daniel, "Targeting In-Kind Transfers Through Market Design: A Revealed Preference Analysis of Public Housing Allocation," *American Economic Review*, 2021, *111* (8), 2660–2696.

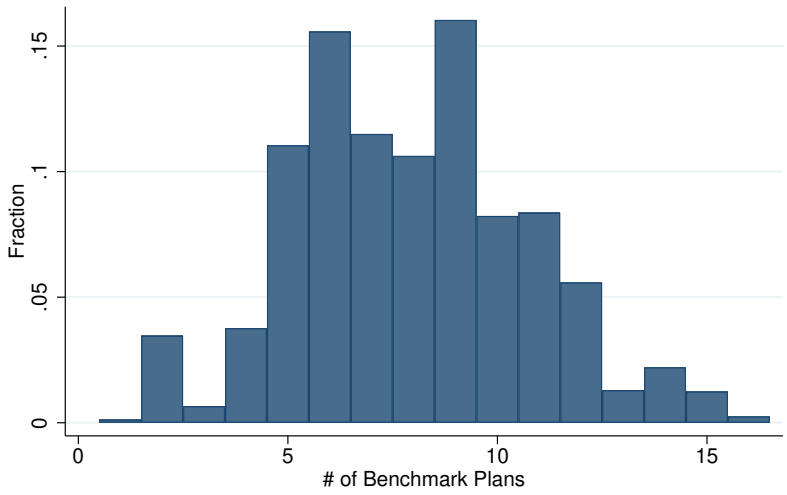
Wallace, Zachary S., Tyler Harkness, Xiaoqing Fu, John H. Stone, Hyon K. Choi, and Rochelle P. Walensky, "Treatment Delays Associated With Prior Authorization for Infusible Medications: A Cohort Study," *Arthritis Care & Research*, 2020, *72* (11), 1543–1549.

Woolhandler, Steffie, Terry Campbell, and David U. Himmelstein, "Costs of Health Care Administration in the United States and Canada," *New England Journal of Medicine*, 2003, *349* (8), 768–775.

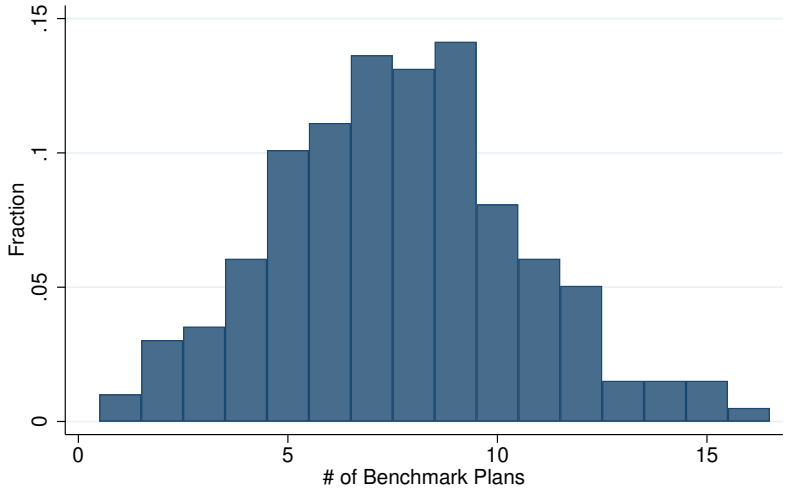
Additional Tables and Figures

Appendix Figure A1. : Distribution of Number of Benchmark Plans in Market-Year

(a) Weighted by Number of Beneficiaries



(b) Unweighted



Notes: This set of figures plots the distribution in the number of benchmark plans across the combination of Part D market region-years. The top figure presents this distribution weighing all Part D market region-years equally, while the bottom weighs Part D market region-years by the number of beneficiaries in our sample enrolled under each.

Appendix Table A1—: Prior authorization use for top drug classes by Medicare Part D spending

	Spending per person-year (USD)	% spending with prior auth	% fills with prior auth
Relatively High Prior Auth			
Tranq/Antipsychotic	185	6.9	3.6
Antivirals	120	14.6	2.1
Antidiabetic Agents, Misc	110	15.0	5.7
Antineoplastic Agents	99	57.7	13.9
CNS Agents, Misc	94	17.6	6.9
Biological Response Modifiers	94	69.6	68.1
Antidepressants	93	7.7	3.3
Cardiac Drugs	88	12.4	5.9
Immunosuppressants	65	66.3	54.7
Anticonvulsants, Misc	60	4.4	1.6
Misc Therapeutic Agents	58	15.0	4.0
Anticoagulants	47	14.5	2.8
NSAIDs	37	10.0	1.6
Vasodilating Agents	27	44.6	1.5
Bone Resorption Inhibitors	22	9.0	4.8
Relatively Low Prior Auth			
Antihyperlipidemic Drugs	212	2.7	1.1
Antidiabetic Agents, Insulins	158	0.6	0.9
Gastrointestinal Drug, Misc	132	2.8	3.2
Opiate Agonists	92	3.5	0.7
Adrenals & Comb	86	3.0	11.6
Antiplatelet Agents	70	0.6	1.4
Cardiac, Calcium Channel	49	1.5	1.0
Anticholinergic	47	0.1	0.2
Cardiac, Beta Blockers	45	0.5	0.5
Eye/Ear/Nose/Throat Misc	44	1.2	0.6
Parasympathomimetic	42	3.2	1.5
Muscle Relaxants	36	1.9	2.3
Antiinflam Agents EENT	29	0.1	0.2
Sympathomimetic Agents	27	2.1	3.4
Estrogens & Comb	25	1.2	5.4

Appendix Table A2—: Estimates of Prior Authorization Request Rejection Rates

Study	Setting	Services	Estimate
Raper et al. (2010)	One HIV clinic, all payers	All drugs	33%
Initial application[a]			
Schwartz et al. (2021)	Large private insurer	Hosp. services and drugs	4.2%
Birdsall et al. (2020)	Academic health system	All drugs	
Initial application			15%
Final application			7.4%
Wallace et al. (2020)	Rheumatology clinic	Infusable drugs	
Initial application			21%
Final application			4%
Kahan et al. (2006)	Israeli MCO	Cefuroxime	8.5%
LaPensee (2003)	One Medicaid MCO	All drugs	4.4%
		Non-formulary drugs	3.7%
		Formulary drugs	7.1%
OEI (2018)	All Medicare Advantage MCOs	All services and drugs	4.1%
AthenaHealth	Physician clients	All drugs	1.5%

Note: This table presents estimates from the literature on the rejection rates associated with requests made for services and drugs restricted under prior authorization. All studies are in U.S. settings unless otherwise noted.
[a] This study does not report final application approval rates in a way that maps on to what we have here.

Appendix Table A3—: First stage regressions restricted to existing users

	$\text{Auth}^{\text{Enrolled}}$	$\text{Excluded}^{\text{Enrolled}}$
$\text{Auth}^{\text{Assigned}}$	0.971	< 0.001
	(< 0.001)	(< 0.001)
$\text{Excluded}^{\text{Assigned}}$	< 0.001	0.966
	(< 0.001)	(< 0.001)
F-statistic	3,610,765	3,482,462
R^2	0.964	0.970
Number of drug × plan × plan years	532,791	
Number of market years	152	
Number of drug × bene years	988,506,911	
Average plans per market year	6.7	
Average benes per plan	627	

Note: This table presents estimates from a set of regressions of prior authorization and exclusion status of drug d in the plan enrolled in by beneficiary i on the status of the plan that they were assigned to. The sample of beneficiaries is restricted only to those who took the drug in the prior year

Appendix Table A4—: Balance Tests

	Spending	# Fills	# Days Supply	% Ever Filled
$\text{Auth}^{\text{Assigned}}$	-0.011	0.000	0.003	0.002
	(0.032)	(0.000)	(0.005)	(0.005)
Reweighted Control Mean	2.651	0.135	0.403	0.307

Note: This table presents estimates from regressions of utilization measures in year $t-1$ on prior authorization and exclusion of a given drug and other drugs within the class in year t. Each observation is a drug-plan-year.

Appendix Table A5—: Main Regressions with alternate utilization variables

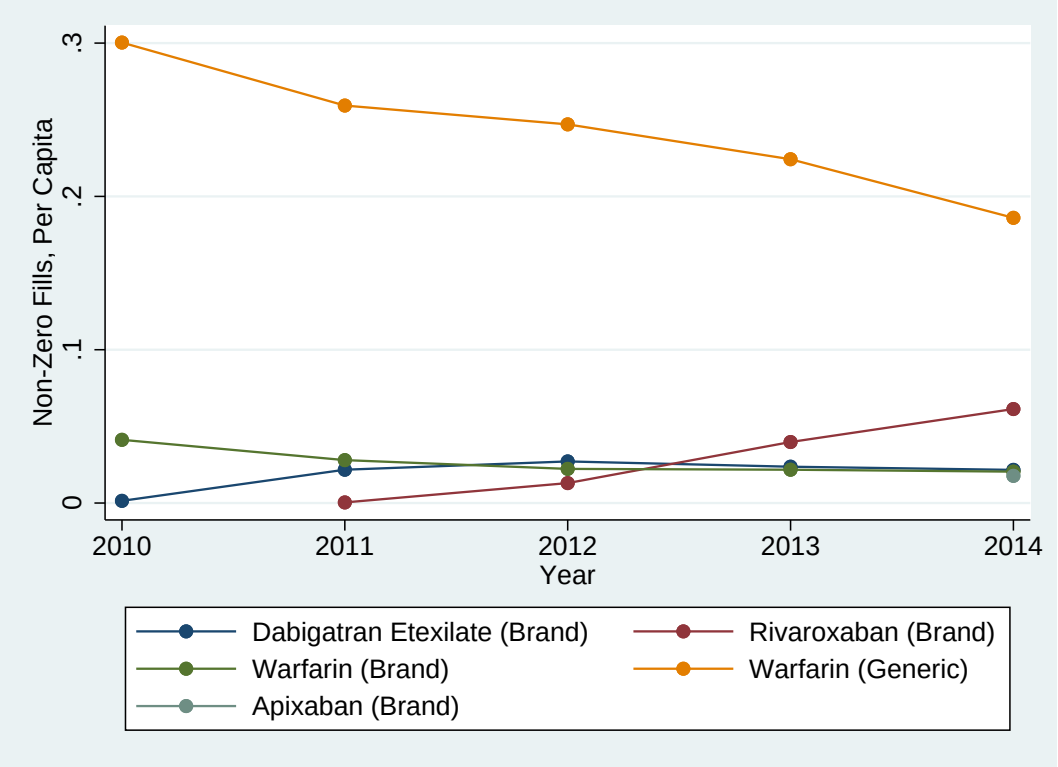
	Days supply per bene		Number of prescriptions per 100 benes per 100 benes	
$\text{Auth}^{\text{Assigned}}$	-0.149	-0.154	-0.455	-0.471
	(0.006)	(0.007)	(0.019)	(0.019)
$\text{Excluded}^{\text{Assigned}}$	-0.206	-0.210	-0.658	-0.672
	(0.0064)	(0.007)	(0.0179)	(0.0179)
Auth^{sub}		0.068		0.207
		(0.007)		(0.023)
$\text{Excluded}^{\text{sub}}$		0.212		0.616
		(0.012)		(0.034)
F-statistic	600	302	788	396
R^2	0.969	0.969	0.966	0.966
Control Mean	1.529		5.067	
Reweighted Control Mean	0.504		1.691	
Number of drug × plan years	2,141,993			
Number of market years	210			
Number of drug × bene years	1,732,564,415			
Average plans per market-yr	6.6			
Average benes per plan	807			

Note: This table presents estimates from regressions of utilization measures on prior authorization and exclusion of a given drug and other drugs within the class. Each observation is a drug-plan-year. Prior authorization of substitute drugs is mean prior authorization status of all other drugs within the class, where drugs are weighted by their average expenditure across all plans in the sample.

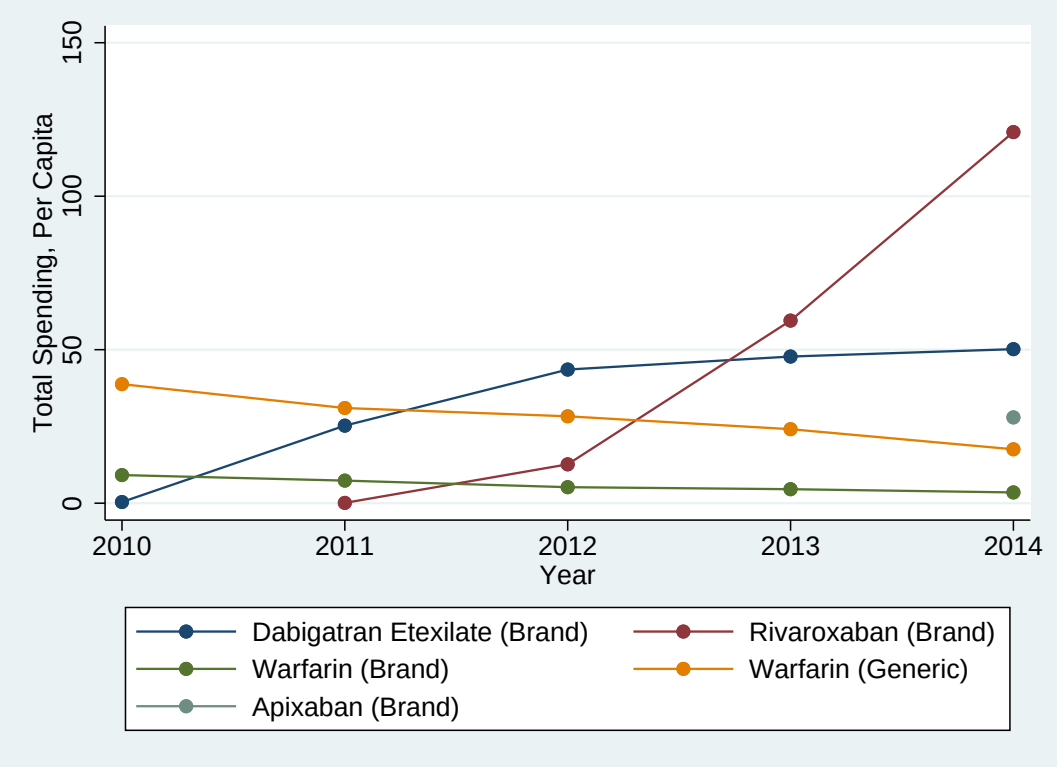
Appendix Table A6—: Main Regressions without plan-region-year fixed effects

	Spending		% Ever filled	
$\text{Auth}^{\text{Assigned}}$	-0.748	-0.767	-0.089	-0.098
	(0.032)	(0.037)	(0.003)	(0.003)
$\text{Excluded}^{\text{Assigned}}$	-0.901	-0.968	-0.110	-0.123
	(0.0269)	(0.032)	(0.0026)	(0.0026)
Auth^{sub}		0.155		0.049
		(0.029)		(0.004)
$\text{Excluded}^{\text{sub}}$		0.519		0.139
		(0.041)		(0.008)
F-statistic	588	269	978	420
R^2	0.831	0.831	0.978	0.978
Control Mean	3.555		1.307	
Reweighted Control Mean	3.613		0.403	
Number of drug × plan years	2,141,993			
Number of market years	210			
Number of drug × bene years	1,732,564,415			
Average plans per market-yr	6.6			
Average benes per plan	807			

Note: This table presents estimates from regressions of utilization measures on prior authorization and exclusion of a given drug and other drugs within the class. Each observation is a drug-plan-year. Prior authorization of substitute drugs is mean prior authorization status of all other drugs within the class, where drugs are weighted by their average expenditure across all plans in the sample.



Appendix Figure A2. : This figure presents the share of patients filling each of these oral anticoagulants at least once during the year, across time.



Appendix Figure A3. : This figure presents the per-patient yearly spending on each of these oral anticoagulants, across time.

Appendix: Prior Authorization Form Examples

Hawki ia health link Amerigroup An Anthem Company

https://providers.amerigroup.com

Novel Oral Anticoagulants Prior Authorization of Benefits Form

CONTAINS CONFIDENTIAL PATIENT INFORMATION
Complete form in its entirety and fax to: Prior Authorization of Benefits Center at 1-844-512-9004.
Provider Help Desk: 1-800-454-3730

1. Patient information	**2. Physician information**
Patient name:_____	Prescribing physician:_____
Patient ID #: _____	Physician address: _____
Patient DOB:_____	Physician phone #: _____
Date of Rx: _____	Physician fax #:_____
Patient phone #:_____	Physician specialty: _____
Patient email address: _____	Physician DEA: _____
	Physician NPI #: _____
	Physician email address: _____

3. Medication	**4. Strength**	**5. Directions**	**6. Quantity per 30 days**
_____	_____	_____	Specify: _____

7. Diagnosis:_____

8. Approval criteria: (Check all boxes that apply. Note: Any areas not filled out are considered not applicable to your patient and may affect the outcome of this request.)

Prior authorization (PA) is not required for preferred novel oral anticoagulants (NOACs). PA is required for nonpreferred NOACs. Requests for doses outside of the manufacturer recommended dose will not be considered. Payment will be considered for FDA approved or compendia indications under the following conditions:

1. Patient does not have a mechanical heart valve.
2. Patient does not have active bleeding.
3. For a diagnosis of atrial fibrillation or stroke prevention, patient has the presence of at least 1 additional risk factor for stroke, with a CHA_2DS_2-VASc score ≥1.
4. A recent creatinine clearance (CrCl) is provided.
5. A recent Child-Pugh score is provided.
6. Patient's current body weight is provided.
7. Patient has documentation of a trial and therapy failure at a therapeutic dose with at least two preferred NOACs.
8. For requests for edoxaban, documentation patient has had 5 to 10 days of initial therapy with a parenteral anticoagulant (low molecular weight heparin or unfractionated heparin). The required trials may be overridden when documented evidence is provided that the use of these agents would be medically contraindicated.

Preferred (no PA required if within established quantity limits)
☐ Eliquis ☐ Xarelto
☐ Pradaxa

Nonpreferred
☐ Savaysa

IAPEC-X1664-19 December 2019

OptumRx has partnered with CoverMyMeds to receive prior authorization requests, saving you time and often delivering real-time determinations.
Visit go.covermymeds.com/OptumRx to begin using this free service.
Please note: All information below is required to process this request.
Mon-Fri: 5am to 10pm Pacific / Sat: 6am to 3pm Pacific

Zetia® (ezetimibe) Prior Authorization Request Form

DO NOT COPY FOR FUTURE USE. FORMS ARE UPDATED FREQUENTLY AND MAY BE BARCODED

Member Information (required)			**Provider Information** (required)		
Member Name:			Provider Name:		
Insurance ID#:			NPI#:	Specialty:	
Date of Birth:			Office Phone:		
Street Address:			Office Fax:		
City:	State:	Zip:	Office Street Address:		
Phone:			City:	State:	Zip:

Medication Information (required)		
Medication Name:	Strength:	Dosage Form:
❑ Check if requesting **brand**	Directions for Use:	
❑ Check if request is for **continuation of therapy**		

Clinical Information (required)

Select the diagnosis below:
- ❑ Homozygous Familial Hypercholesterolemia (HoFH)
- ❑ Homozygous Sitosterolemia
- ❑ Primary Hypercholesterolemia
- ❑ Other diagnosis: _____ ICD-10 Code(s):_____

Clinical information:
Has the patient's diagnosis been confirmed? ❑ **Yes** ❑ **No**

Select the medications the patient has a failure, contraindication, or intolerance to:
- ❑ Ezetimibe-simvastatin
- ❑ Lovastatin
- ❑ Simvastatin
- ❑ Other statin or statin combination product. Please specify all: _____

Quantity limit requests:
What is the quantity requested per DAY? _____

What is the reason for exceeding the plan limitations?
- ❑ Titration or loading dose purposes
- ❑ Patient is on a dose-alternating schedule (e.g., one tablet in the morning and two tablets at night, one to two tablets at bedtime)
- ❑ Requested strength/dose is not commercially available
- ❑ Other: _____

Are there any other comments, diagnoses, symptoms, medications tried or failed, and/or any other information the physician feels is important to this review?

_____

_____

Please note: This request may be denied unless all required information is received.
For urgent or expedited requests please call 1-800-711-4555.
This form may be used for non-urgent requests and faxed to 1-800-527-0531.

ANTIPSYCHOTICS
PRIOR AUTHORIZATION FORM
(form effective 1/5/21)

Keystone First | PERFORMRX℠ Next Generation Pharmacy Benefits

Fax to PerformRx℠ at **1-215-937-5018**, or to speak to a representative call **1-800-588-6767**.

PRIOR AUTHORIZATION REQUEST INFORMATION

☐ New request ☐ Renewal request	Total pages:	Office contact/phone:	LTC facility contact/phone:

PATIENT INFORMATION

Patient name:		Patient ID#:	DOB:
Street address:	Apt #:	City/state/zip:	

PRESCRIBER INFORMATION

Prescriber name:			
Specialty:		NPI:	State license #:
Street address:	Suite #:	City/state/zip:	
Phone:	Fax:		

MEDICATION REQUESTED

Preferred Agents

☐ Abilify Maintena	☐ fluphenazine elixir	☐ haloperidol tablet	☐ Invega Sustenna	☐ Perseris ER injection	☐ risperidone tablet
☐ aripiprazole tablet	☐ fluphenazine oral concentrate	☐ haloperidol decanoate inj	☐ Invega Trinza	☐ quetiapine tablet	☐ trifluoperazine tablet
☐ Aristada ER injection	☐ fluphenazine tablet	☐ haloperidol lactate inj.	☐ loxapine capsule	☐ quetiapine ER tablet	☐ ziprasidone capsule
☐ Aristada Initio injection	☐ fluphenazine decan. inj.	☐ haloperidol lactate oral concentrate	☐ olanzapine tablet	☐ Risperdal Consta	☐ Zyprexa Relprevv
☐ clozapine tablet	☐ Haldol injection		☐ perphenazine tablet	☐ risperidone solution	

Non-Preferred Agents

☐ Abilify Mycite	☐ chlorpromazine tablet	☐ Geodon injection	☐ olanzapine inj/ODT	☐ Saphris SL tablet	☐ Versacloz suspension
☐ Abilify tablet	☐ clozapine ODT	☐ Haldol decanoate inj.	☐ olanzapine/fluoxetine cap	☐ Secuado patch	☐ Vraylar capsule
☐ Adasuve inhalation	☐ Clozaril tablet	☐ Invega ER tablet	☐ paliperidone ER tab	☐ Seroquel tablet	☐ Zyprexa tablet/injection
☐ amitriptyline/perphenazine	☐ Fanapt tablet	☐ Latuda tablet	☐ pimozide tablet	☐ Seroquel XR tablet	☐ Zyprexa Zydis
☐ aripiprazole ODT	☐ Fazaclo dispersible tablet	☐ molindone tablet	☐ Rexulti tablet	☐ Symbyax capsule	☐ other:
☐ aripiprazole solution	☐ fluphenazine HCl injection	☐ Nuplazid capsule	☐ Risperdal solution/tablet	☐ thioridazine tablet	
☐ Caplyta capsules	☐ Geodon capsule	☐ Nuplazid tablet	☐ risperidone ODT	☐ thiothixene capsule	

Strength:	Dosage form:	Directions:	Quantity:	Refills:
Diagnosis:			Diagnosis code (required):	

PHARMACY INFORMATION (Prescriber to identify the pharmacy that is to dispense the medication):

Deliver to: ☐ Patient's Home ☐ Physician's Office ☐ Patient's Preferred Pharmacy Name:	
Pharmacy Phone #:	Pharmacy Fax #:
☐ I acknowledge that the patient agrees with the pharmacy chosen for delivery of this medication.	

REQUEST FOR A NON-PREFERRED AGENT

1. Has the patient taken the requested non-preferred antipsychotic in the past 90 days?
   ☐ Yes – *Submit documentation.* ☐ No
2. Has the patient tried and failed the preferred medications (listed above)?
   ☐ Yes – List medications tried: ☐ No
3. Does the patient have a contraindication or intolerance to the preferred medications?
   ☐ Yes – *Submit documentation of contraindication/intolerance.* ☐ No
4. **For oral Invega/paliperidone ER requests, does the patient have active liver disease with elevated LFTs or is the patient at risk for active liver disease?**
   ☐ Yes – *Submit documentation and lab values.* ☐ No

REQUEST FOR A PATIENT LESS THAN 18 YEARS OF AGE

5. Is this request for a dose increase of a previously approved medication? ☐ Yes – *Submit recent chart documentation supporting the increased dose.* ☐ No
6. Is the requested agent prescribed by, or in consultation with, one of the following physician specialists? ☐ Yes ☐ No *Submit documentation of consultation, if applicable.*
   ☐ child development pediatrician ☐ child & adolescent psychiatrist ☐ general psychiatrist (only if patient is ≥ 14 years of age) ☐ pediatric neurologist
7. Does the patient have severe behavioral problems related to a psychotic or neuro-developmental disorder? ☐ Yes – *Submit medical record documentation.* ☐ No
8. Has the patient tried non-drug therapies? ☐ Yes – *Submit medical record documentation.* ☐ No
9. Has the patient had the following baseline and/or follow-up monitoring? Check all that apply. ☐ BMI (or weight/height) ☐ blood pressure ☐ fasting glucose level ☐ fasting lipid panel
   ☐ presence of extrapyramidal symptoms (EPS) using the Abnormal Involuntary Movement Scale (AIMS)
   Submit documentation of all monitoring/test results.

REQUEST FOR A LOW-DOSE ORAL ANTIPSYCHOTIC FOR A PATIENT 18 YEARS OF AGE OR OLDER

10. What is the TOTAL daily dose of the requested medication? _____mg/day
    Submit documentation of complete medication regimen.
11. Is the low dose prescribed as part of a plan to titrate up to a therapeutic dose? ☐ Yes – *Submit documentation of titration plan.* ☐ No

REQUEST FOR THERAPEUTIC DUPLICATION OF AN ATYPICAL OR TYPICAL ANTIPSYCHOTIC

12. Does the patient have a medical reason for concomitant use of the requested medications? ☐ Yes – *Submit documentation with justification.* ☐ No
13. Is this request for a drug that is being titrated to, or tapered from, a drug in the same class? ☐ Yes – *List medication.* ☐ No

PLEASE FAX COMPLETED FORM WITH REQUIRED CLINICAL DOCUMENTATION

Prescriber signature:	Date: